ENGINEERING THE HUMAN VITAMIN D RECEPTOR TO BIND A NOVEL SMALL MOLECULE: INVESTIGATING THE STRUCTURE-FUNCTION RELATIONSHIP BETWEEN HUMAN VITAMIN D RECEPTOR AND VARIOUS LIGANDS

A Thesis
Presented to
The Academic Faculty

by

Amanda Ousley

In Partial Fulfillment
of the Requirements for the Degree
Doctor of Philosophy in the
School of Chemistry and Biochemistry

Georgia Institute of Technology
May 2011

ENGINEERING THE HUMAN VITAMIN D RECEPTOR TO BIND A NOVEL SMALL MOLECULE: INVESTIGATING THE STRUCTURE-FUNCTION RELATIONSHIP BETWEEN HUMAN VITAMIN D RECEPTOR AND VARIOUS LIGANDS

Approved by:

Dr. Donald Doyle, Advisor
School of Chemistry and Biochemistry
Georgia Institute of Technology

Dr. Nicholas Hud
School of Chemistry and Biochemistry
Georgia Institute of Technology

Dr. Bahareh Azizi, Advisor
School of Chemistry and Biochemistry
Georgia Institute of Technology

Dr. Wendy L. Kelly
School of Chemistry and Biochemistry
Georgia Institute of Technology

Dr. Andreas Bommarius, Co-Advisor
School of Chemical and Biomolecular Engineering
Georgia Institute of Technology

Dr. Loren Williams, Co-Advisor
School of Chemistry and Biochemistry
Georgia Institute of Technology

Date Approved: April 11, 2011

This work is dedicated to my mother.
For all of your love and support

ACKNOWLEDGEMENTS

I would like to begin by thanking all of my committee members for their help and support throughout my time here. I would like to thank Dr. Chiaolong Hsiao for his help with learning how to use the ITC and interpret the data. I would like to thank the Hud lab for the use of the ITC. I would next like to thank the Tolbert lab for the use of the fluorometer for my fluorescence experiments. I would like to thank the Ortland lab at Emory for their generous donation of the pMalRXR plasmid as well as the TEV protease plasmid, without which my experiments would not have been completed. I would like to thank Janna Blum for her assistance with protein purification. I would also like to thank her for running my HPLC experiments. I would like to thank Naman Kanakiya, my undergraduate, for assisting in the making of most of the mutants that were needed. Without his help the mutagenesis would not have been possible. Next, I would like to thank Anna Duraj-Thatte for all of her help and expertise in protein expression and purification. I would like to thank Michael Rood, for all of his scientific input on experiements and papers. Next, I would like to thank Hilda Castillo and Jennifer Taylor. We have been through everything together and come out on the other side better for this experience. I would not have wanted to go through this with anyone but them. I have to give extreme thanks to Hally Shaffer. Hally is a great sounding board for research and for life. Without her these past years would have been a lot more difficult than they already were. I would like to thank my advisor, Dr. Donald Doyle, for starting me on my path to this point. I would also like to thank the past members of the Doyle lab for their guidance and support. Finally, I would like to thank Dr. Bahareh Azizi. Her support and

guidance through these past years has been indescribable. I can say for sure that I would not be in this position right now if not for her pushing me to succeed and be better than I even thought I could be. I could not have done this without her.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF SYMBOLS AND ABBREVIATIONS

$1,25(OH)_2D_3$	1α,25-Dihydroxyvitamin D_3
3-AT	3-Amino 1,2 Triazole
AF-1	Activation Function 1
AF-2	Activation Function 2
CMV	Cytomegalovirus
CoAC	Coactivator
DBD	DNA Binding Domain
ER	Estrogen Receptor
GAD	Gal4 Activation Domain
GBD	Gal4 DNA Binding Domain
HAT	Histone Acetyl Transferase
HDAC	Histone Deacetylase
HEK293T	Human Embryonic Kidney Cells
HPLC	High Performance Liquid Chromatography
HRE	Hormone Response Element
hVDR	Human Vitamin D Receptor
LBD	Ligand Binding Domain
LBP	Ligand Binding Pocket
LCA	Lithocholic Acid
MBP	Maltose Binding Protein
NCoR	Nuclear Receptor Corepressor
NR	Nuclear Receptor
PXR	Pregnane X Receptor

RAR	Retinoic Acid Receptor
RXR	Retinoid X Receptor
SC-ALW	Adenine Selective Media
SMRT	Silencing Mediator for Retinoid and Thyroid Hormone Receptors
TEV	Tobacco Etch Virus
UV	Ultraviolet

SUMMARY

The human vitamin D receptor (hVDR) is a member of the nuclear receptor superfamily, involved in calcium and phosphate homeostasis; hence implicated in a number of diseases, such as rickets and osteoporosis. This receptor binds 1α,25-dihydroxyvitamin D_3 (also referred to as $1,25(OH)_2D_3$) and other known ligands, such as lithocholic acid. Specific interactions between the receptor and ligand are crucial for the function and activation of this receptor, as implied by the single point mutation, H305Q, causing symptoms of Type II rickets. In this work, further understanding of the significant and essential interactions between the ligand and the receptor were deciphered, through a combination of rational and random mutagenesis. A hVDR mutant, H305F, was engineered with increased sensitivity towards lithocholic acid, with an EC_{50} value of 10 μM and 40 $\pm$ 14 fold activation in mammalian cell assays, while maintaining wild-type activity with $1,25(OH)_2D_3$. Furthermore, via random mutagenesis, H305F/H397Y was discovered to bind cholecalciferol, a precursor in the 1α,25-dihydroxyvitamin D_3 biosynthetic pathway, which does not activate wild-type hVDR. The variant H305F/H397Y is activated by cholecalciferol concentrations as low as 100 nM, with an EC_{50} value of 300 nM and 70 $\pm$ 11 fold activation in mammalian cell assays. *In silico* docking analysis of the variant displays a dramatic conformational shift of cholecalciferol in the ligand binding pocket in comparison to the docked analysis of cholecalciferol with wild-type hVDR. This shift is hypothesized to be due to the introduction of two bulkier residues, suggesting that the addition of these bulkier residues

introduces molecular interactions between the ligand and receptor, leading to activation with cholecalciferol.

CHAPTER 1
NUCLEAR RECEPTORS

1.1 Nuclear Receptor Superfamily

Nuclear receptors (NRs) are ligand-activated transcription factors. Upon ligand binding, these receptors activate the transcription of essential genes required for a number of cellular and physiological processes including proliferation, differentiation, and development [1-4]. For example, the human vitamin D receptor controls the transcription of genes involved in calcium and phosphate homeostasis [5]. Also, the pregnane X receptor is involved in upregulating the expression of genes that are involved in detoxification of the body [6]. Due to their involvement in gene expression throughout mammalian systems, NRs are implicated in a number of diseases including, cancer, diabetes, and osteoporosis [5, 7-10].

To date, there are 48 known human nuclear receptors [8]. NRs fall into three subfamilies: steroid receptors, such as estrogen receptor; non-steroidal receptors, such as retinoid X receptor; and orphan receptors, such as liver receptor homolog-1 [7, 8, 11-14]. Steroid receptor ligands include estrogen and testoterone; whereas non-steroidal receptors can bind steroid-like ligands or other molecules, such as fatty acids or lipids [2, 8, 13]. The orphan receptors are a subfamily of nuclear receptors whose wild-type ligand is unknown, such as steroidogenic factor-1 [15]. These receptors bind a wide variety of ligands due to the differences in their ligand binding pockets [8].

1.2 Nuclear Receptor Structure

NRs consist of five to six domains, labeled A to F: a N-terminal domain (A/B), a DNA binding domain (DBD) (C), a hinge region (D), a ligand binding domain (LBD)

(E), and a C-terminal domain (F) (Figure 1.1) [1, 2, 11, 16, 17]. Each of these domains plays a role in the structure and function of NRs. The A/B region consists of the N-terminal portion of the protein and contains the activation function-1 (AF-1) domain. This domain is highly disordered and the sequence of this domain varies, lacking conservation among the members of the superfamily [7, 18, 19]. The AF-1 domain is capable of functioning independently. For example, transcriptional activation has been shown to occur with the estrogen receptor AF-1 domain in the presence of tamoxifen [7, 18-20]. The AF-1 domain has also been shown to be the target of post-translational modifications and to interact indirectly with transcriptional machinery [20, 21]. Similar to the A/B domain, the F domain is also not well characterized, but this domain has been implicated in ligand binding and nuclear receptor dimerization [11].

The DBD binds to short specific sequences of DNA known as hormone response elements (HRE), and is the most conserved domain among NRs, sharing approximately 95% sequence identity among members of the superfamily [1, 4, 7, 8, 13, 19]. Crystallographic studies have shown that the DBD is predominately α-helical and consists of two cysteine-rich zinc finger motifs. Each zinc finger contains four cysteines coordinated to one zinc ion, allowing the α-helices to bind to specific base pairs in the major groove of the DNA [11, 22, 23]. The DBD is divided into subdomains known as the P and D boxes (Figure 1.2) [24]. These regions have been shown to determine HRE specificity, define the DBD dimerization interface, and make contacts with the flanking regions of DNA surrounding the HREs [11, 25]. The P and D boxes determine the HRE

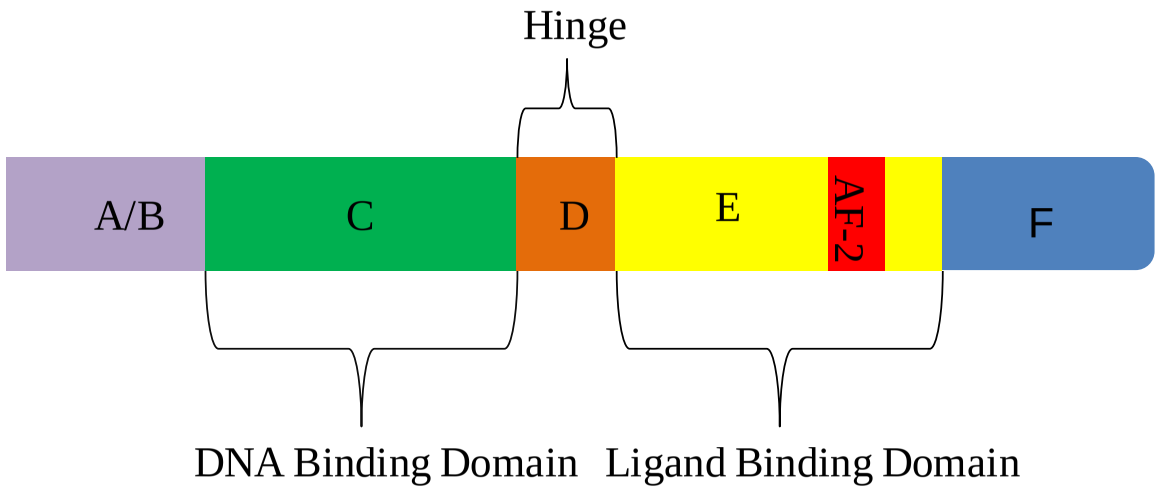

Figure 1.1: Nuclear Receptor Domains

specificity and spacing, respectively [11, 26, 27]. HREs can be located upstream or downstream of the target gene, and are oriented as half-sites (direct-, indirect-, or inverted repeats). Direct repeats are homologous sequences of DNA that are separated by a number of spacing base pairs; these base pairs are specific to each NR monomer, homodimer, or heterodimer (i.e. AGTTCAtgagAGTTCA). Inverted repeats contain complementary sequences separated by a spacer (i.e. GACTGCtgacGCAGTC). Due to the fact that most NRs function as homo- or heterodimers, a half-site is required for each partner. The majority of steroid receptors function as homodimers, which bind to inverted repeats, whereas most non-steroidal receptors function as heterodimers and bind to direct repeats [11, 28, 29].

The hinge region, or D domain, conjoins the DBD and LBD. This domain allows for rotation of the DBD, reducing steric hindrance between the two domains [11]. The D domain contains a nuclear localization sequence, which is necessary for the translocation of NRs into the nucleus where transcription occurs. Some nuclear receptors, such as the steroid receptors, bind a ligand in the cytosol, and are then transported into the nucleus to interact with the corresponding HRE. NRs such as the glucocorticoid receptor, androgen receptor, estrogen receptor, and progesterone receptor bind ligands prior to translocation [1, 7, 11, 30]. Conversely, NRs such as the retinoid X receptor, liver X receptor, pregnane X receptor, and the human vitamin D receptor translocate into the nucleus without bound ligand [2, 5, 8, 31].

The LBD plays a role in dimerization, ligand binding, and coregulator interactions [19]. The majority of NR LBDs consists of 12 α-helices (H1-H12) and a short β-turn, forming three layers of antiparallel α-helices (Figure 1.3) [2, 30, 32]. In most NRs, H1-

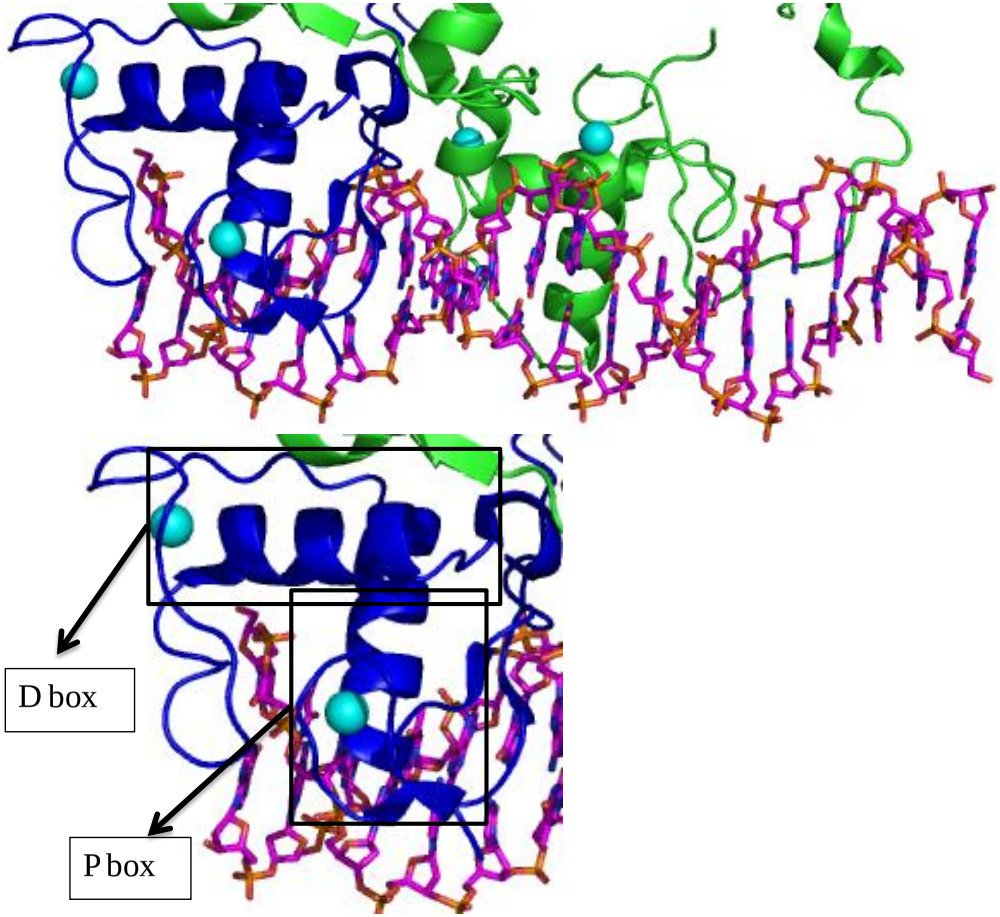

Figure 1.2: DBD bound to DNA (PDB:3DZY). Retinoid X receptor DBD= Blue, Peroxisome proliferator-activated receptor γ DBD= Green, Zinc ions= Cyan, and DNA = magenta.

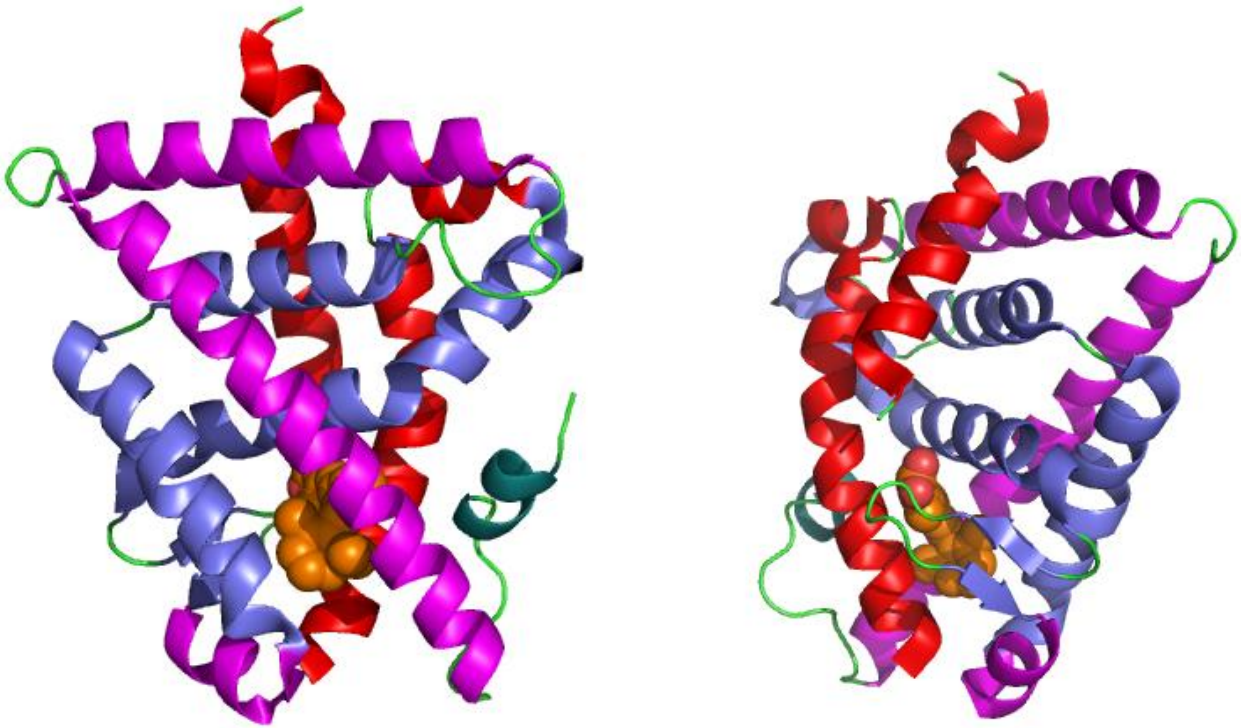

Figure 1.3: Two different views of RXR LBD (PDB:1FBY). A front view (left) as well as a side view (right) are shown of the LBD. Most nuclear receptors consist of 12 α-helices and 2 β-strands. The helices and β-sheet create three layers. Layer1, which is made up of H1-H3 is shown in red. The central layer which consists of H4, H5, β-turn, H8, and H9 is slate blue and Layer 3 which consists of H6, H7, and H10 is shown in magenta. Helix 12 is shown in teal, and the ligand (9-*cis* retinoic acid) is shown in orange.

H3 comprise one of the layers, whereas the H4, H5, the β-turn, H8, and H9 make up the central layer of the LBD. H6, H7, and H10 make up the third layer of the LBD. Upon ligand binding, H3, H5, and H12 undergo a conformational change [19, 30, 32-34]. One of the most critical helices, H12, consists of the activation function-2 (AF-2) domain, which plays a critical role in the ligand activation of NRs, through interactions with coactivators [19, 30].

While nuclear receptor LBDs share varying degrees of similarity with their overall secondary structure content, the differences in their ligand binding pockets (LBP) account for the diverse ligand binding profiles of these nuclear receptors [8, 11, 30, 35]. LBP's vary based on size, shape, volume, and chemical properties among various nuclear receptors [11, 36]. For example, the pregnane X receptor (PXR) has a pocket volume of ~1200 $Å^3$, while the estrogen receptor (ER) has a pocket volume of ~450 $Å^3$ [13, 31]. The chemical properties or polarity of the LBP also contribute in determining the different types of ligands capable of binding the receptors. Polar or charged residues can provide hydrogen bonding interactions between the ligand and the receptor, as observed with the hydrogen bonds formed between the 3-hydroxyl group of the "A" ring of estradiol and arginine 394 and glutamate 353 of ER. These hydrogen bonds are crucial for stabilizing the ligand in the pocket, leading to activation of the receptor. The polar and charged residues can also form salt bridges between surrounding residues and the ligand, as can be observed with the carboxylate of all-*trans* retinoic acid and arginine 278 and serine 289 of retinoic acid receptor (RAR) [37]. Hydrophobic residues can provide van der Waals interactions between the ligand and receptor, such as that seen with

isoleucine 268, alanine 272, and tryptophan 305 of retinoid X receptor, which are known to form hydrophobic contacts with 9-*cis* retinoic acid [1, 2, 8, 11, 12, 30].

1.3 Nuclear Receptor Function

As ligand activated transcription factors, NRs function in a precise manner involving a series of molecular events that lead to regulation of essential genes [1, 30, 38]. NRs generally function as homodimers or as heterodimers with the retinoid X receptor (RXR) [1, 11, 39, 40]. Heterodimers are further classified into three categories: permissive, conditional, and non-permissive [39]. Permissive partners require the respective ligands of both heterodimer partners in order for activation of transcription to occur. An example is seen with the RXR and the pregnane X receptor (PXR), in which case transcriptional activation requires both ligands, 9-*cis* retinoic acid and rifampicin (synthetic/non-natural ligand), respectively. Conditional partners, such as RXR and retinoic acid receptor, require only one of the ligands to be present for activation; however, maximal activation is achieved if both ligands are present [39]. Finally, non-permissive partners only require the ligand of the NR partner and not the heterodimerization partner RXR, as seen in the case of vitamin D receptor (VDR) and retinoid X receptor, where only 1α,25-dihydroxyvitamin D_3 is needed for activation of transcription [39].

As ligand activated transcription factors, NRs are involved in gene repression and activation. These ligands vary in size, shape, and chemical properties, such as rifampicin, 27-hydroxycholesterol, and estradiol [2, 13, 41, 42]. Most nuclear receptors must bind a ligand (agonist) for transcription to occur and very few are considered to be constitutively active, or activated without the presence of a bound ligand. NRs are also capable of binding other ligands (antagonists or inverse agonists) that lead to repression of

transcription [1, 11]. Most antagonists function by displacing the agonist from the ligand binding pocket, while inverse agonists function by binding to receptors that express constitutively or have high basal activity, repressing transcriptional activation [43-45].

Repression of transcription via NRs can occur one of three ways. In the structure of an apo NR, the AF-2 domain is positioned away from the receptor, exposing the corepressor binding site on the LBD, allowing corepressors to bind. Common human corepressors include the nuclear receptor corepressor (NCoR) and silencing mediator for retinoid and thyroid hormone receptors (SMRT), each of which binds to a specific hydrophobic groove on the NR LBD, through the CoR box that contains a LXXXIXXXI/L motif [46-50]. The binding of the corepressor induces the recruitment of transcriptional complexes, including histone deacetylases (HDACs) [51]. HDACs deacetylate lysine residues in the N-terminal tail of histone complexes, leading to a more tightly packed chromatin structure and preventing RNA polymerase from initiating transcription (Figure 1.4A) [43, 44, 51, 52].

Agonist binding causes structural changes throughout the receptor, particularly in H3, H5, and H12. The most notable change occurs in H12, where the AF-2 domain is rearranged, positioning H12 closer to the LBD, and serving as a seal to the LBP (Figure 1.5) [53]. The conformational change leads to the creation of a hydrophobic groove between H3, H5, and H12, for coactivator (CoAc) binding [54-57]. CoAcs contain multiple LXXLL motifs, collectively known as the LXXLL NR box, which bind to the nuclear receptor in the hydrophobic groove. The binding of these proteins leads to the recruitment of large complexes that include histone acetyltransferases (HATs) [58-60].

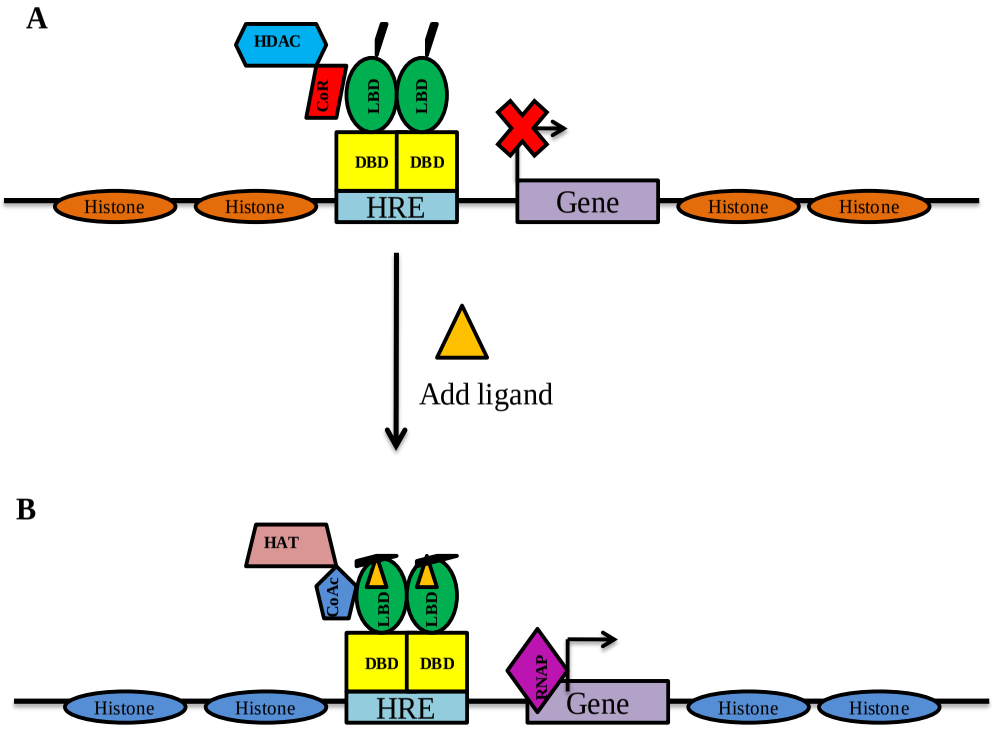

Figure 1.4: Nuclear Receptor Transcriptional Repression and Activation Scheme. (A) Transcriptional repression occurs when the nuclear receptor is associated with corepressors. These proteins can recruit histone deacetylases (HDACs), which can deacetylate the histones and represss transcription. (B) Transcriptional activation occurs when ligand binds to the nuclear receptor, recruiting coactivators and histone acetyl transferases (HATs). HATs acetylate the histones, allowing RNA polymerase to access the DNA and transcription to occur.

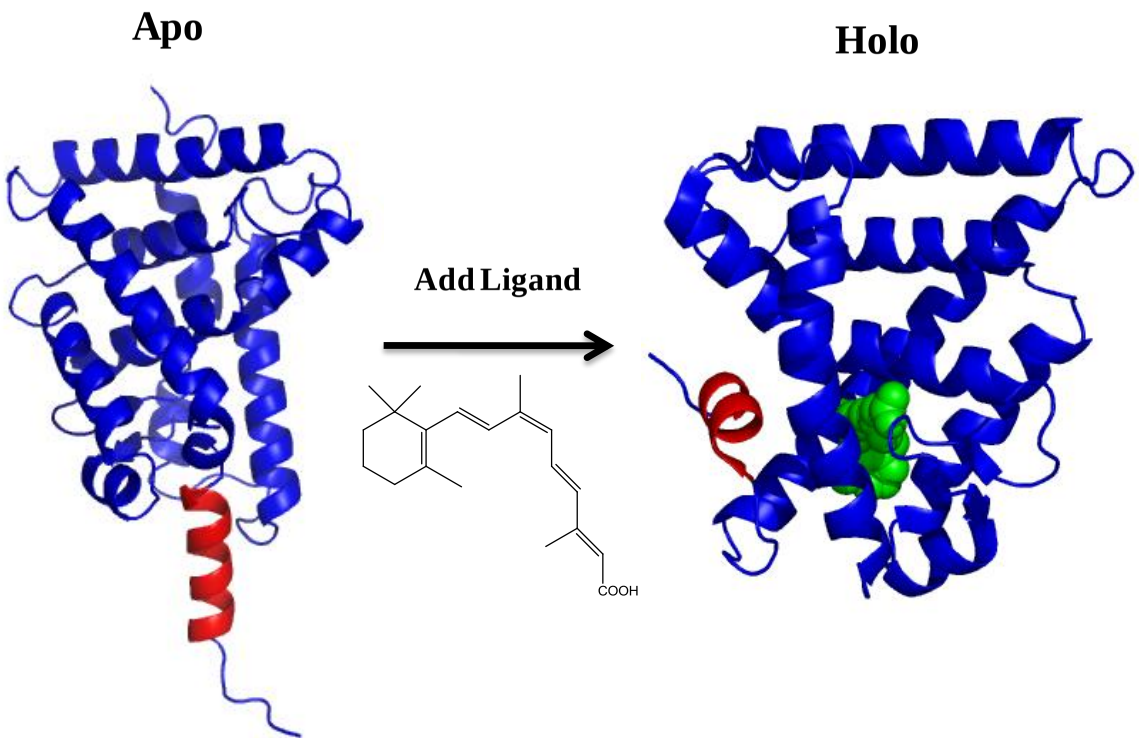

Figure 1.5: Apo- and Holo- Nuclear Receptor. Apo- represents that no ligand is bound, whereas Holo- means that ligand is bound. Red= Helix 12 and Green= 9-*cis* retinoic acid (ligand)

HATs acetylate the lysine residues on the N-terminal tail of histones leading to a less compact histone complex, allowing RNA polymerase to access the DNA and initiate transcription (Figure 1.4B) [55, 58-60].

Due to their implications in a number of diseases and their involvement in transcriptional activation and repression, the development of new drugs for nuclear receptor based diseases is of vast interest to pharmaceutical companies. To date, approximately 10% of the most commonly prescribed drugs on the market are NR ligands [7, 8]. Some of these drugs include: peroxisome proliferator-activated receptor ligands, such as thiazolidiones, used to treat Type II Diabetes. One of the most commonly known drugs that binds a nuclear receptor is tamoxifen, an estrogen receptor antagonist, used to treat breast cancer [9, 10]. To discover new drugs for NRs, understanding of the structure/function relationship between nuclear receptors and their various ligands is critical.

1.4 Human Vitamin D Receptor

The human vitamin D receptor (hVDR) is a nuclear receptor primarily involved in biological processes such as apoptosis, immune responses, and calcium and phosphate homeostasis [61]. hVDR was first cloned from the human intestine in 1988 and since that time, 11 orthologs have been discovered [62-68]. Each VDR ortholog binds the same natural ligand, 1α,25-dihydroxyvitamin D_3 ($1,25(OH)_2D_3$) [62-69]. In addition to $1,25(OH)_2D_3$, VDR is known to bind a variety of $1,25(OH)_2D_3$ analogs and bile acids, such as lithocholic acid (LCA), which displays an EC_{50} value of 12 μM in comparison to an EC_{50} value of 1 nM for hVDR and $1,25(OH)_2D_3$ (Figure 1.6) [12, 70-74].

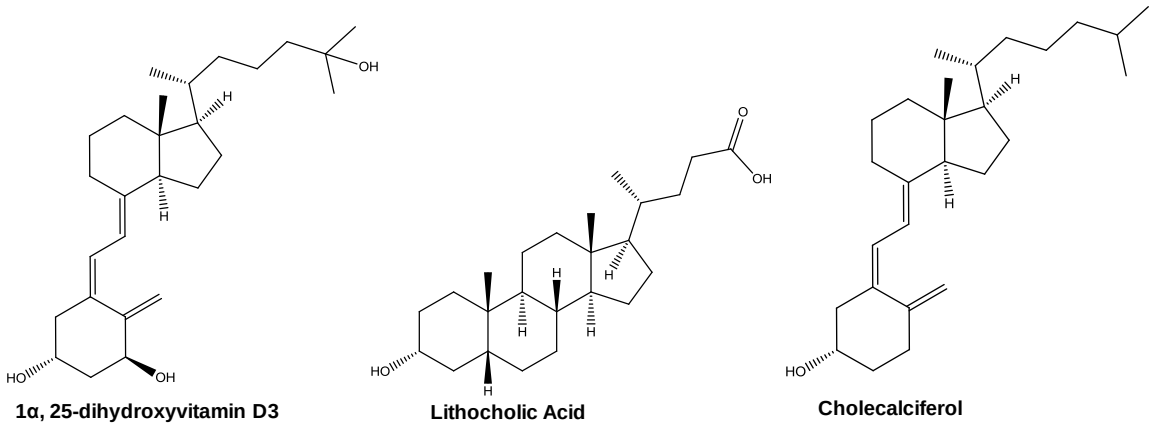

Figure 1.6: hVDR ligands. 1α,25-Dihydroxyvitamin D_3 and lithocholic acid are hVDR ligands. Cholecalciferol is an inactive precursor to 1α,25-dihydroxyvitamin D_3.

hVDR is a nuclear receptor comprised of 427 amino acids [62]. The ligand binding domain of this receptor is comprised of 303 amino acids with a ligand binding pocket volume of 693 $Å^3$. The LBP is elongated and consists of polar and charged residues at the ends of the pocket, and hydrophobic residues surrounding the pocket interior (Figure 1.7). These residues contribute essential molecular interactions between the ligand and receptor, such as hydrogen bonding and van der Waals forces, which are required for receptor activation and function. Disruption of these key interactions reduces the activity of a functional receptor and in some cases, leads to detrimental effects. For example, a single point mutation in the LBD of hVDR, H305Q, causes a 10-fold decrease in sensitivity of the receptor to $1,25(OH)_2D_3$, causing Type II rickets [75, 76]. Vitamin D analogs used to treat hVDR-related diseases have higher calcemic effects than $1,25(OH)_2D_3$. To design new drugs for the treatment of these diseases, a better

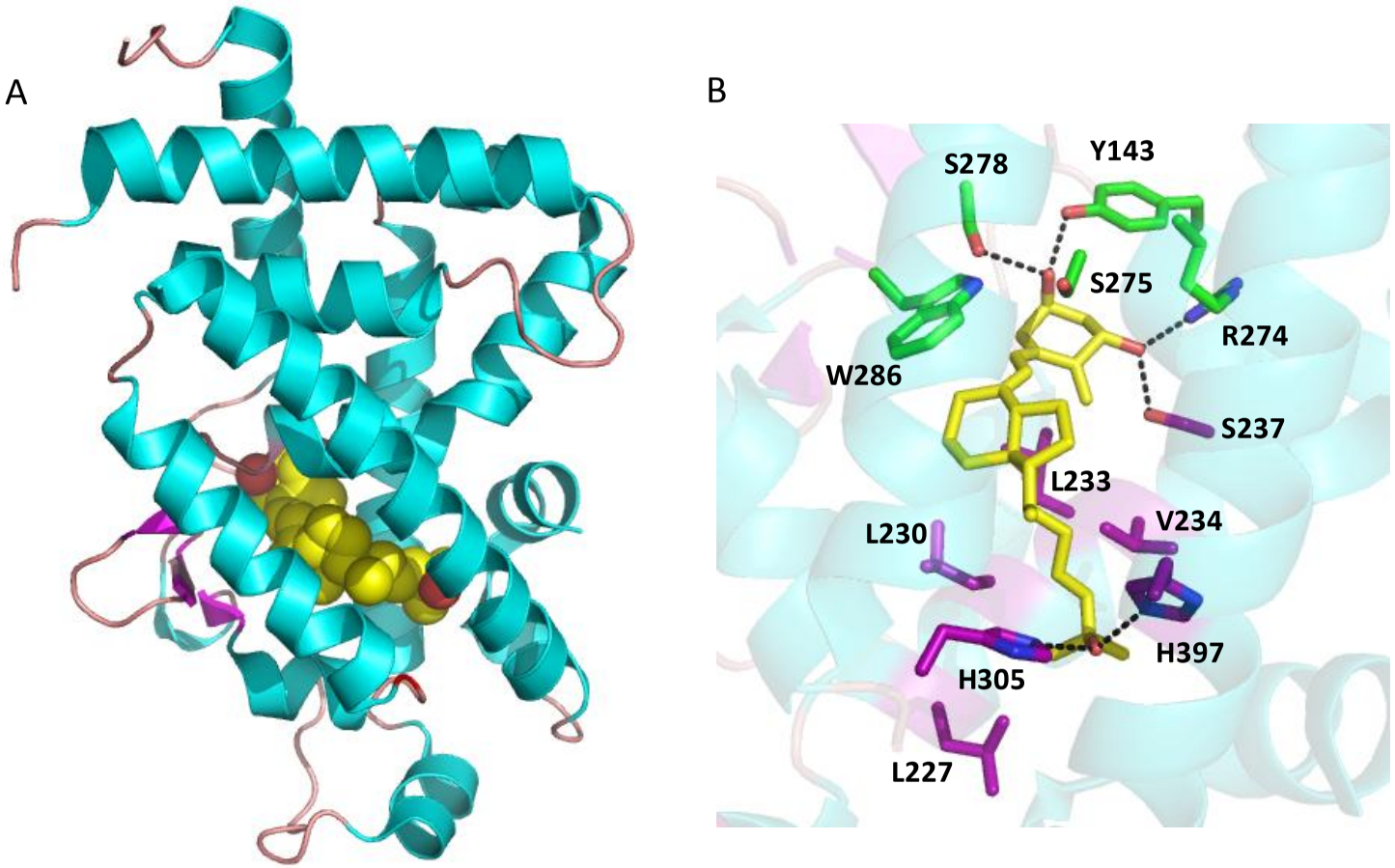

Figure 1.7: hVDR crystal structure. (A) hVDR LBD crystal structure bound to 1α,25-dihydroxyvitaminD_3. (B) hVDR LBP with important residues shown.
(- - - - - - -) are known hydrogen bonding residues.

understand the structure/function relationship between hVDR and various ligands is needed.

hVDR is implicated in several diseases, such as osteoporosis, hypocalcemia, hyperthyroidism, diabetes, and rickets [5]. Rickets (Type I and Type II), is a disorder stemming from a vitamin D deficiency and impared hVDR function [77-79]. This results in growth retardation, bone deformities, hypocalcemia, and secondary hyperthyroidism [77, 80]. In Type II rickets, mutations in the ligand binding domain of hVDR prevent the ability of the natural ligand, $1,25(OH)_2D_3$, from binding to the receptor [77]. However, in Type I rickets, the 1α-hydroxylase, responsible for the hydroxylation of 25-hydroxyvitamin D to its active form, $1,25(OH)_2D_3$, is nonfunctional. Thus, patients with Type I rickets are unable to generate active vitamin D and are treated with $1,25(OH)_2D_3$ supplements [77, 79, 80]. There is great interest in the design of vitamin D analogs able to bind hVDR, which counteract the vitamin D deficiency caused by Type II rickets [81]. $1,25(OH)_2D_3$ is rapidly broken down by the body; therefore, synthetic analogs, which are not broken down as rapidly, are being designed. These analogs need to also be designed with fewer calcemic effects so that their prolonged presence in the body reduces negative effects on the pathways in which hVDR is involved [5]. The goal of this work is to engineer a hVDR variant that is capable of binding a novel small molecule, cholecalciferol. Through this process, a better understanding of the structure/function relationship between hVDR and various ligands will be attained, which will aid in designing new hVDR ligands.

1.5 Summary

Due to the ability of NRs to bind a variety of ligands, their involvement in a number of biological processes, and their implication in a number of diseases, nuclear receptors make great targets for drug design and protein engineering. Protein engineering of nuclear receptors will allow us to develop a better understanding of structure/function relationship between the receptors and ligands. hVDR makes a good target for protein engineering due the lack of structure/function information known about this receptor. Also, the knowledge gained via protein engineering will allow for a better understanding of this receptor such as which residues are necessary for ligand binding and activation.

1.6 References

1. Mangelsdorf DJ, Thummel C, Beato M, Herrlich P, Schutz G, Umesono K, Blumberg B, Kastner P, Mark M, Chambon P *et al*: **The nuclear receptor superfamily - the second decade** *Cell* 1995, **83**(6):835-839.

2. Egea PF, Mitschler A, Rochel N, Ruff M, Chambon P, Moras D: **Crystal structure of the human RXR alpha ligand-binding domain bound to its natural ligand: 9-cis retinoic acid**. *Embo J* 2000, **19**(11):2592-2601.

3. Khorasanizadeh S, Rastinejad F: **Nuclear-receptor interactions on DNA-response elements**. *EMBO J* 2001, **26**(6):384-390.

4. Chambon P: **The nuclear receptor superfamily: a personal retrospect on the first two decades**. *Mol Endocrinol* 2005, **19**(6):1418-1428.

5. Jones G, Strugnell SA, DeLuca HF: **Current understanding of the molecular actions of vitamin D**. *Physiol Rev* 1998, **78**(4):1193-1231.

6. Kliewer SA, Goodwin B, Willson TM: **The nuclear pregnane X receptor: a key regulator of xenobiotic metabolism**. *Endocr Rev* 2002, **23**(5):687-702.

7. Gronemeyer H, Gustafsson JA, Laudet V: **Principles for modulation of the nuclear receptor superfamily**. *Nat Rev Drug Discov* 2004, **3**(11):950-964.

8. Ottow E, Weinmann H (eds.): **Nuclear receptors as drug targets**. Weinheim: Wiley-VCH Verlag GmbH & Co. KGaA; 2008.

9. Knight WA, Livingston RB, Gregory EJ, McGuire WL: **Estrogen receptor as an independent prognostic factor for early recurrence in breast cancer**. *Cancer Res* 1977, **37**(12):4669-4671.

10. Spiegelman BM: **PPAR-gamma: adipogenic regulator and thiazolidinedione receptor**. *Diabetes* 1998, **47**(4):507-514.

11. Germain P, Staels B, Dacquet C, Spedding M, Laudet V: **Overview of nomenclature of nuclear receptors**. *Pharmacol Rev* 2006, **58**(4):685-704.

12. Rochel N, Wurtz JM, Mitschler A, Klaholz B, Moras D: **The crystal structure of the nuclear receptor for vitamin D bound to its natural ligand**. *Mol Cell* 2000, **5**(1):173-179.

13. Brzozowski AM, Pike AC, Dauter Z, Hubbard RE, Bonn T, Engstrom O, Ohman L, Greene GL, Gustafsson JA, Carlquist M: **Molecular basis of agonism and antagonism in the oestrogen receptor**. *Nature* 1997, **389**(6652):753-758.

14. Fayard E, Auwerx J, Schoonjans K: **LRH-1: an orphan nuclear receptor involved in development, metabolism and steroidogenesis**. *Trends Cell Biol* 2004, **14**(5):250-260.

15. Jameson JL: **Of mice and men: The tale of steroidogenic factor-1**. *J Clin Endocrinol Metab* 2004, **89**(12):5927-5929.

16. Giguere V, McBroom LD, Flock G: **Determinants of target gene specificity for ROR alpha 1: monomeric DNA binding by an orphan nuclear receptor**. *Mol Cell Biol* 1995, **15**(5):2517-2526.

17. Krust A, Green S, Argos P, Kumar V, Walter P, Bornert JM, Chambon P: **The chicken oestrogen receptor sequence: homology with v-erbA and the human oestrogen and glucocorticoid receptors**. *EMBO J* 1986, **5**(5):891-897.

18. Folkertsma S, van Noort P, Van Durme J, Joosten HJ, Bettler E, Fleuren W, Oliveira L, Horn F, de Vlieg J, Vriend G: **A family-based approach reveals the function of residues in the nuclear receptor ligand-binding domain**. *J Mol Biol* 2004, **341**(2):321-335.

19. Warnmark A, Treuter E, Wright AP, Gustafsson JA: **Activation functions 1 and 2 of nuclear receptors: molecular strategies for transcriptional activation**. *Mol Endocrinol* 2003, **17**(10):1901-1909.

20. Metivier R, Penot G, Flouriot G, Pakdel F: **Synergism between ERalpha transactivation function 1 (AF-1) and AF-2 mediated by steroid receptor coactivator protein-1: requirement for the AF-1 alpha-helical core and for a direct interaction between the N- and C-terminal domains**. *Mol Endocrinol* 2001, **15**(11):1953-1970.

21. Lavery DN, McEwan IJ: **Structure and function of steroid receptor AF1 transactivation domains: induction of active conformations**. *Biochem J* 2005, **391**(Pt 3):449-464.

22. Luisi BF, Xu WX, Otwinowski Z, Freedman LP, Yamamoto KR, Sigler PB: **Crystallographic analysis of the interaction of the glucocorticoid receptor with DNA**. *Nature* 1991, **352**(6335):497-505.

23. Schwabe JW, Chapman L, Finch JT, Rhodes D: **The crystal structure of the estrogen receptor DNA-binding domain bound to DNA: how receptors discriminate between their response elements**. *Cell* 1993, **75**(3):567-578.

24. Chandra V, Huang P, Hamuro Y, Raghuram S, Wang Y, Burris TP, Rastinejad F: **Structure of the intact PPAR-gamma-RXR- nuclear receptor complex on DNA**. *Nature* 2008, **456**(7220):350-356.

25. Umesono K, Evans RM: **Determinants of target gene specificity for steroid/thyroid hormone receptors**. *Cell* 1989, **57**(7):1139-1146.

26. Laudet V: **Evolution of the nuclear receptor superfamily: early diversification from an ancestral orphan receptor**. *J Mol Endocrinol* 1997, **19**(3):207-226.

27. Zilliacus J, Carlstedt-Duke J, Gustafsson JA, Wright AP: **Evolution of distinct DNA-binding specificities within the nuclear receptor family of transcription factors**. *Proc Natl Acad Sci U S A* 1994, **91**(10):4175-4179.

28. Beato M, Herrlich P, Schutz G: **Steroid hormone receptors: many actors in search of a plot**. *Cell* 1995, **83**(6):851-857.

29. Rastinejad F, Perlmann T, Evans RM, Sigler PB: **Structural determinants of nuclear receptor assembly on DNA direct repeats**. *Nature* 1995, **375**(6528):203-211.

30. Moras D, Gronemeyer H: **The nuclear receptor ligand-binding domain: structure and function**. *Curr Opin Cell Biol* 1998, **10**(3):384-391.

31. Watkins RE, Wisely GB, Moore LB, Collins JL, Lambert MH, Williams SP, Willson TM, Kliewer SA, Redinbo MR: **The human nuclear xenobiotic receptor PXR: structural determinants of directed promiscuity**. *Science* 2001, **292**(5525):2329-2333.

32. Benoit G, Malewicz M, Perlmann T: **Digging deep into the pockets of orphan nuclear receptors: insights from structural studies**. *Trends Cell Biol* 2004, **14**(7):369-376.

33. Chambon P: **A decade of molecular biology of retinoic acid receptors**. *FASEB J* 1996, **10**(9):940-954.

34. Steinmetz ACU, Renaud JP, Moras D: **Binding of Ligands and Activation of Transcription by Nuclear Receptors**. *Annu Rev Biophys Biomolec Struct* 2001, **30**:329-359.

35. Schwimmer LJ, Rohatgi P, Azizi B, Seley KL, Doyle DF: **Creation and discovery of ligand-receptor pairs for transcriptional control with small molecules**. *Proc Natl Acad Sci U S A* 2004, **101**(41):14707-14712.

36. Germain P, Kammerer S, Perez E, Peluso-Iltis C, Tortolani D, Zusi FC, Starrett J, Lapointe P, Daris JP, Marinier A *et al*: **Rational design of RAR-selective ligands revealed by RARbeta crystal stucture**. *EMBO Rep* 2004, **5**(9):877-882.

37. Renaud JP, Rochel N, Ruff M, Vivat V, Chambon P, Gronemeyer H, Moras D: **Crystal structure of the RAR-gamma ligand-binding domain bound to all-trans retinoic acid**. *Nature* 1995, **378**(6558):681-689.

38. Zhang ZD, Burch PE, Cooney AJ, Lanz RB, Pereira FA, Wu JQ, Gibbs RA, Weinstock G, Wheeler DA: **Genomic analysis of the nuclear receptor family: new insights into structure, regulation, and evolution from the rat genome**. *Genome Res* 2004, **14**(4):580-590.

39. Shulman AI, Larson C, Mangelsdorf DJ, Ranganathan R: **Structural determinants of allosteric ligand activation in RXR heterodimers**. *Cell* 2004, **116**(3):417-429.

40. Mangelsdorf DJ, Evans RM: **The RXR heterodimers and orphan receptors**. *Cell* 1995, **83**(6):841-850.

41. Chrencik JE, Orans J, Moore LB, Xue Y, Peng L, Collins JL, Wisely GB, Lambert MH, Kliewer SA, Redinbo MR: **Structural disorder in the complex of**

human pregnane X receptor and the macrolide antibiotic rifampicin. *Mol Endocrinol* 2005, **19**(5):1125-1134.

42. Janowski BA, Willy PJ, Devi TR, Falck JR, Mangelsdorf DJ: **An oxysterol signalling pathway mediated by the nuclear receptor LXR alpha**. *Nature* 1996, **383**(6602):728-731.

43. Love JD, Gooch JT, Nagy L, Chatterjee VK, Schwabe JW: **Transcriptional repression by nuclear receptors: mechanisms and role in disease**. *Biochem Soc Trans* 2000, **28**(4):390-396.

44. Nagy L, Kao HY, Love JD, Li C, Banayo E, Gooch JT, Krishna V, Chatterjee K, Evans RM, Schwabe JW: **Mechanism of corepressor binding and release from nuclear hormone receptors**. *Genes Dev* 1999, **13**(24):3209-3216.

45. Pike AC, Brzozowski AM, Walton J, Hubbard RE, Bonn T, Gustafsson JA, Carlquist M: **Structural aspects of agonism and antagonism in the oestrogen receptor**. *Biochem Soc Trans* 2000, **28**(4):396-400.

46. Chen JD, Evans RM: **A transcriptional co-repressor that interacts with nuclear hormone receptors**. *Nature* 1995, **377**(6548):454-457.

47. Horlein AJ, Naar AM, Heinzel T, Torchia J, Gloss B, Kurokawa R, Ryan A, Kamei Y, Soderstrom M, Glass CK *et al*: **Ligand-independent repression by the thyroid hormone receptor mediated by a nuclear receptor co-repressor**. *Nature* 1995, **377**(6548):397-404.

48. Ordentlich P, Downes M, Xie W, Genin A, Spinner NB, Evans RM: **Unique forms of human and mouse nuclear receptor corepressor SMRT**. *Proc Natl Acad Sci U S A* 1999, **96**(6):2639-2644.

49. Park EJ, Schroen DJ, Yang M, Li H, Li L, Chen JD: **SMRTe, a silencing mediator for retinoid and thyroid hormone receptors-extended isoform that is more related to the nuclear receptor corepressor**. *Proc Natl Acad Sci U S A* 1999, **96**(7):3519-3524.

50. Zamir I, Harding HP, Atkins GB, Horlein A, Glass CK, Rosenfeld MG, Lazar MA: **A nuclear hormone receptor corepressor mediates transcriptional silencing by receptors with distinct repression domains**. *Mol Cell Biol* 1996, **16**(10):5458-5465.

51. Guenther MG, Barak O, Lazar MA: **The SMRT and N-CoR corepressors are activating cofactors for histone deacetylase 3**. *Mol Cell Biol* 2001, **21**(18):6091-6101.

52. Nagy L, Kao HY, Chakravarti D, Lin RJ, Hassig CA, Ayer DE, Schreiber SL, Evans RM: **Nuclear receptor repression mediated by a complex containing SMRT, mSin3A, and histone deacetylase**. *Cell* 1997, **89**(3):373-380.

53. Bourguet W, Germain P, Gronemeyer H: **Nuclear receptor ligand-binding domains: three-dimensional structures, molecular interactions and pharmacological implications**. *Trends Pharmacol Sci* 2000, **21**(10):381-388.

54. Chen JD: **Steroid/nuclear receptor coactivators**. *Vitam Horm* 2000, **58**:391-448.

55. Vo N, Goodman RH: **CREB-binding protein and p300 in transcriptional regulation**. *J Biol Chem* 2001, **276**(17):13505-13508.

56. McKenna NJ, O'Malley BW: **Combinatorial control of gene expression by nuclear receptors and coregulators**. *Cell* 2002, **108**(4):465-474.

57. Chen D, Huang SM, Stallcup MR: **Synergistic, p160 coactivator-dependent enhancement of estrogen receptor function by CARM1 and p300**. *J Biol Chem* 2000, **275**(52):40810-40816.

58. Belandia B, Parker MG: **Nuclear receptors: a rendezvous for chromatin remodeling factors**. *Cell* 2003, **114**(3):277-280.

59. Leo C, Chen JD: **The SRC family of nuclear receptor coactivators**. *Gene* 2000, **245**(1):1-11.

60. Stallcup MR, Kim JH, Teyssier C, Lee YH, Ma H, Chen D: **The roles of protein-protein interactions and protein methylation in transcriptional activation by nuclear receptors and their coactivators**. *J Steroid Biochem Mol Biol* 2003, **85**(2-5):139-145.

61. Hourai S, Rodrigues LC, Antony P, Reina-San-Martin B, Ciesielski F, Magnier BC, Schoonjans K, Mourino A, Rochel N, Moras D: **Structure-based design of a superagonist ligand for the vitamin D nuclear receptor**. *Chem Biol* 2008, **15**(4):383-392.

62. Baker AR, McDonnell DP, Hughes M, Crisp TM, Mangelsdorf DJ, Haussler MR, Pike JW, Shine J, Omalley BW: **Cloning and expression of full-length cDNA-**

encoding human vitamin-D receptor. *Proc Natl Acad Sci U S A* 1988, **85**(10):3294-3298.

63. Jehan F, DeLuca HF: **Cloning and characterization of the mouse vitamin D receptor promoter**. *Proc Natl Acad Sci U S A* 1997, **94**(19):10138-10143.

64. Whitfield GK, Dang HT, Schluter SF, Bernstein RM, Bunag T, Manzon LA, Hsieh G, Dominguez CE, Youson JH, Haussler MR *et al*: **Cloning of a functional vitamin D receptor from the lamprey (Petromyzon marinus), an ancient vertebrate lacking a calcified skeleton and teeth**. *Endocrinology* 2003, **144**(6):2704-2716.

65. Craig TA, Sommer S, Sussman CR, Grande JP, Kumar R: **Expression and regulation of the vitamin D receptor in the zebrafish, Danio rerio**. *J Bone Miner Res* 2008, **23**(9):1486-1496.

66. Lu Z, Hanson K, DeLuca HF: **Cloning and origin of the two forms of chicken vitamin D receptor**. *Arch Biochem Biophys* 1997, **339**(1):99-106.

67. Chun RF, Chen H, Boldrick L, Sweet C, Adams JS: **Cloning, sequencing, and functional characterization of the vitamin D receptor in vitamin D-resistant New World primates**. *Am J Primatol* 2001, **54**(2):107-118.

68. Suzuki T, Suzuki N, Srivastava AS, Kurokawa T: **Identification of cDNAs encoding two subtypes of vitamin D receptor in flounder, *Paralichthys olivaceus***. *Biochem Biophys Res Commun* 2000, **270**(1):40-45.

69. Reddy MD, Stoynova L, Acevedo A, Collins ED: **Residues of the human nuclear vitamin D receptor that form hydrogen bonding interactions with the three hydroxyl groups of 1alpha,25-dihydroxyvitamin D3**. *J Steroid Biochem Mol Biol* 2007, **103**(3-5):347-351.

70. Vaisanen S, Ryhanen S, Saarela JTA, Maenpaa PH: **Structure-function studies of new C-20 epimer pairs of vitamin D_3 analogs**. *Eur J Biochem* 1999, **261**(3):706-713.

71. Nayeri S, Kahlen JP, Carlberg C: **The high affinity ligand binding conformation of the nuclear 1,25-dihydroxyvitamin D_3 receptor is functionally linked to the transactivation domain 2 (AF-2)**. *Nucleic Acids Res* 1996, **24**(22):4513-4518.

72. Jurutka PW, Thompson PD, Whitfield GK, Eichhorst KR, Hall N, Dominguez CE, Hsieh JC, Haussler CA, Haussler MR: **Molecular and functional**

comparison of 1,25-dihydroxyvitamin D_3 and the novel vitamin D receptor ligand, lithocholic acid, in activating transcription of cytochrome P450 3A4. *J Cell Biochem* 2005, **94**(5):917-943.

73. Adachi R, Honma Y, Masuno H, Kawana K, Shimomura L, Yamada S, Makishima M: **Selective activation of vitamin D receptor by lithocholic acid acetate, a bile acid derivative**. *J Lipid Res* 2005, **46**(1):46-57.

74. Mizwicki MT, Bula CM, Mahinthichaichan P, Henry HL, Ishizuka S, Norman AW: **On the mechanism underlying (23*S*)-25-dehydro-1alpha(OH)-vitamin D3-26,23-lactone antagonism of hVDRwt gene activation and its switch to a superagonist**. *J Biol Chem* 2009, **284**(52):36292-36301.

75. Malloy PJ, Eccleshall TR, Gross C, VanMaldergem L, Bouillon R, Feldman D: **Hereditary vitamin D resistant rickets caused by a novel mutation in the vitamin D receptor that results in decreased affinity for hormone and cellular hyporesponsiveness**. *J Clin Invest* 1997, **99**(2):297-304.

76. Rochel N, Hourai S, Moras D: **Crystal structure of hereditary vitamin D-resistant rickets--associated mutant H305Q of vitamin D nuclear receptor bound to its natural ligand**. *J Steroid Biochem Mol Biol*, **121**(1-2):84-87.

77. Hughes MR, Malloy PJ, Kieback DG, Kesterson RA, Pike JW, Feldman D, Omalley BW: **Point mutations in the human vitamin-D receptor gene associated with hypocalcemic rickets**. *Science* 1988, **242**(4886):1702-1705.

78. Kitanaka S, Takeyama K, Murayama A, Sato T, Okumura K, Nogami M, Hasegawa Y, Niimi H, Yanagisawa J, Tanaka T *et al*: **Inactivating mutations in the 25-hydroxyvitamin D-3 1 alpha-hydroxylase gene in patients with pseudovitamin D-deficiency rickets**. *N Engl J Med* 1998, **338**(10):653-661.

79. Malloy PJ, Hochberg Z, Tiosano D, Pike JW, Hughes MR, Feldman D: **The Molecular Basis of Hereditary 1,25-Dihydroxyvitamin-D_3 Resistant Rickets in Seven Related Families**. *J Clin Invest* 1990, **86**(6):2071-2079.

80. Kitanaka S, Takeyama K, Murayama A, Kato S: **The Molecular basis of vitamin D-Dependent Rickets Type I**. *Endocr J* 2001, **48**(4):427-432.

81. Fraser D, Kooh SW, Kind HP, Holick MF, Tanaka Y, Deluca HF: **Pathogenesis of hereditary vitamin-D-dependent rickets- inborn error of vitamin-D metabolism involving defective conversion of 25-hydroxyvitamin-D to 1alpha,25-dihydroxyvitamin-D**. *N Engl J Med* 1973, **289**(16):817-822.

CHAPTER 2
ENGINEERING HUMAN VITAMIN D RECEPTOR LIBRARIES

2.1 Introduction

2.1.1 Human Vitamin D Receptor Background

Nuclear receptors are potential targets for protein engineering due to their involvement in a variety of biological processes and their ability to bind a wide variety of ligands. The diversity in the LBP's implies that specific combinations of residues are necessary for binding specific ligands. Therefore, a better understanding of the structure/function relationship between the receptor and ligand is necessary. hVDR was chosen for protein engineering based on this receptors' implication in diseases, as well as the impact of a single point mutation on the receptor function, as seen with the H305Q mutation of hVDR, which leads to Type II rickets [1, 2]. A better understanding of the residues that are important in the binding of various ligands, as well as the mutational tolerance in hVDR, can be used towards developing new drugs that can be used to treat hVDR-related diseases. The development of new drugs is crucial for the treatment of hVDR related diseases. Also, better understanding of the important residues of this receptor will aid in further overall knowledge of the structure and function relationship of this receptor.

Previously, structural and mutational studies with hVDR were performed to assess the role of residues within the ligand binding pocket important for ligand binding and activation [3-12]. Crystal structures of hVDR with $1,25(OH)_2D_3$ and other agonists, such as MC1288 and KH1060, as well as mutational analyses, such as alanine scanning, have provided useful preliminary information regarding interactions vital to ligand binding.

The aforementioned crystal structures maintain the same hydrogen bonding interactions with Y143, S237, R274, S278, H305, and H397. However, the aliphatic chain arrangements of the ligands allow for different contacts with the surrounding residues including: L309, L404, and V418. Hydrogen bonding residues were determined via alanine scanning mutagenesis, and residues necessary for activation were identified by determining which residue mutations lead to receptor inactivation [9, 12, 13].

Initially, Y143, S237, R274, S278, H305, and H397 were found to form essential hydrogen bonds with $1,25(OH)_2D_3$, and all of these residues except S278 are essential to obtain wild-type activation of hVDR with $1,25(OH)_2D_3$ (Figure 2.1) [3, 6-9, 12, 13]. However, Mizwicki *et al.* have since shown that the hydrogen bonds between H305, H397, and the ligand are unnecessary for activation with $1,25(OH)_2D_3$. Mutating C288 to a glycine or alanine and W286 to phenylalanine or alanine was also shown to inhibit transcriptional activation [5, 10, 11, 13]. Previous structural analysis of the hVDR residues and the work of Mizwicki *et al.* has prompted a better understanding regarding the mutational tolerance of these residues. The mutational tolerance of the hVDR LBP will allow for a more thorough understanding about crucial residues in ligand binding and activation as well as determining residues that are capable of being mutated and maintain activation. Three protein libraries were designed and tested using a method called chemical complementation, which links the survival of yeast to the activation of the nuclear receptor by a small molecule.

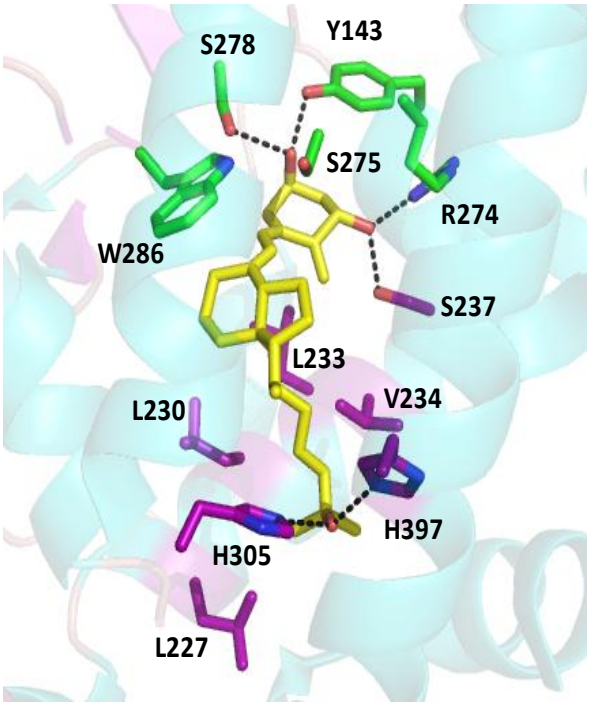

Figure 2.1: hVDR ligand binding pocket. Known important residues involved in ligand activation are shown. For example, Y143 and R274 form hydrogen bonds hydroxyl groups of the ligand, whereas W286 and L233 are known to form hydrophobic contacts with the “C” and “D” rings of the ligand.

2.1.2 Chemical Complementation

Chemical complementation is a three-component system in yeast that links the survival of yeast to the function of a nuclear receptor and a small molecule [14-17]. This system is comprised of three-components: a Gal4 DNA binding domain (GBD) fused to a nuclear receptor LBD, a nuclear receptor coactivator bound to a Gal4 activation domain (GAD), and small molecule. Gal4 is a ligand independent yeast transcription factor, which consists of a DNA binding domain and an activation domain. The yeast strain PJ69-4A is used, which consists of Gal4 response elements controlling the expression of selective markers, such as *ADE2* and *HIS3* [18]. In this system, when the small molecule ligand binds to the nuclear receptor/GBD fusion, a conformational change recruits the coactivator/Gal4 activation domain fusion. Upon the GBD binding to the Gal4 response elements, the GAD recruits the yeast transcriptional machinery, activating transcription of the selective marker and allowing for yeast survival (Figure 2.2). Thus, chemical complementation links the survival of yeast to the binding of a small molecule to the nuclear receptor of interest [14]. There are numerous advantages of using chemical complementation as a system for analyzing NR/ligand pairs. The high transformation efficiency for yeast on the order of 10^6 allows for analysis of large libraries of variants. Yeast are simple eukaryotes, and undergo the same transcriptional processes as mammalian cells. However, their lack of NRs, prevents background activation from NRs as observed in mammalian cells.

Applications of chemical complementation include drug discovery, protein engineering, and metabolic/enzyme engineering. Drug discovery for nuclear receptors is a fast growing field due to NR's implications in disease, and chemical complementation

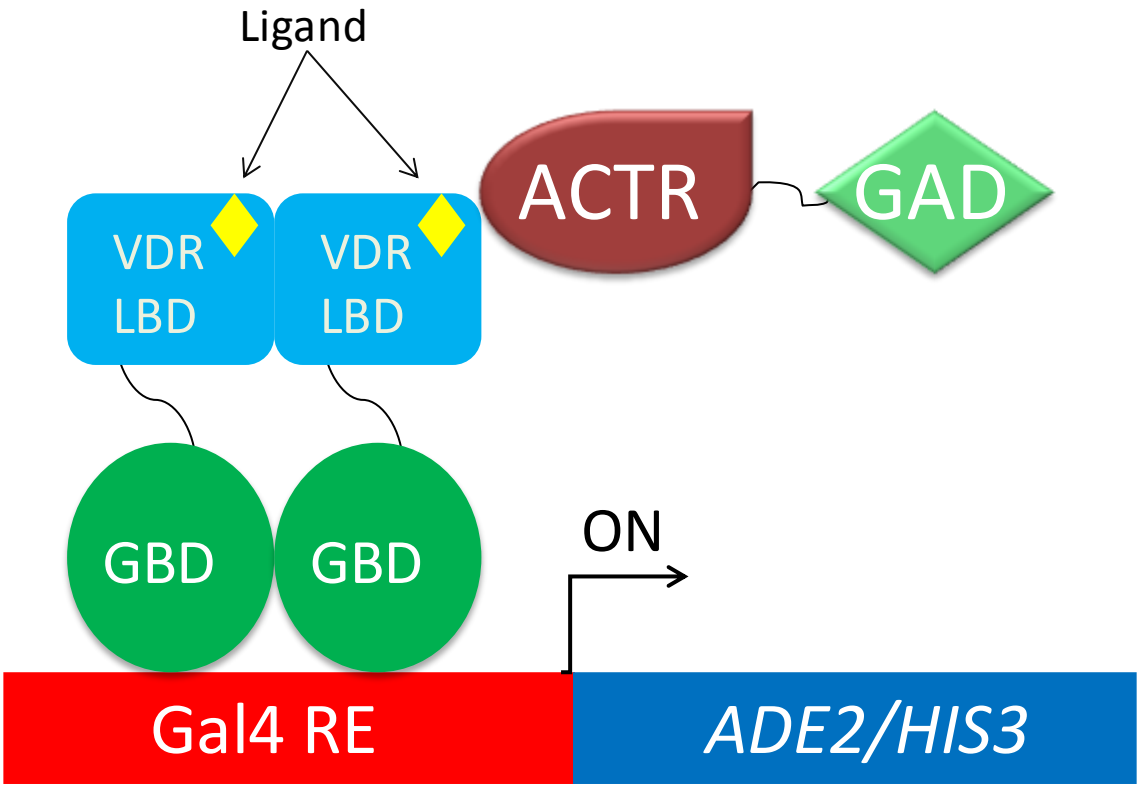

Figure 2.2: Chemical Complementation Schematic. A yeast strain has the Gal4 response elements control the expression of a yeast selective marker (*ADE2/HIS3*). Two fusion proteins are necessary for the system to function. A fusion protein containing the NRLBD fused to a Gal4 DNA binding domain as well as a fusion protein containing a human coactivator fused to the Gal4 activation domain are needed for this system. Upon ligand binding to the LBD the coactivator/GAD fusion will recruit transcriptional machinery and turn on the expression of the selective marker.

can be used to analyze drugs for NR-based diseases in a semi high-throughput method [14, 15]. Using chemical complementation for drug discovery instead of mammalian cells eliminates the background that might be observed from NRs that are present in mammalian cells. Also, a large number of drugs can be screened in a 48 hours as well as the fact that yeast are simple eukaryotes and therefore contain the same transcriptional processes as mammalian cells.

Chemical complementation can also be used for protein engineering, specifically with NRs. In the past, chemical complementation has been used to identify novel nuclear receptor/ligand pairs using a NR variant that has been engineered to bind a synthetic small molecule ligand incapable of binding the wild-type NR. For example, the retinoid X receptor has been engineered to bind a novel small molecule ligand, LG335 instead of the wild-type ligand, 9-*cis* retinoic acid [15, 16]. These engineered receptors can be used as molecular switches in gene therapy. In gene therapy, an exogenous gene is introduced into the body. However, there needs to be a mechanism for the control of expression of the gene. Therefore, the engineered nuclear receptor could be used to turn the expression of the gene on or off by the addition of the synthetic ligand.

Also, the engineered nuclear receptors can be used to detect the production of small molecules from enzyme libraries. This system can be used for evaluation of enzyme libraries in a metabolic pathway for the production of a small molecule by linking the survival of the yeast to the production of the small molecule. In this system a nuclear receptor would be engineered to bind the target product of the enzymatic reaction. If an enzyme in the library can produce the desired product, the product can bind and activate

the NR, causing transcription to occur and the yeast will survive. Therefore, chemical complementation is a valuable method for analyzing hVDR libraries.

2.1.3 Testing hVDR with Chemical Complementation

The activity of hVDR in chemical complementation was determined using a fusion of a hVDR LBD with a GBD (pGBDhVDR) containing a tryptophan marker, and a plasmid containing coactivator/GAD fusion protein (pGAD10BAACTR) and a leucine marker. Both plasmids were transformed into the yeast strain, PJ69-4A. Colonies were selected using medium lacking leucine and tryptophan, selecting for both plasmids containing the fusion proteins. Yeast cells were then added to an adenine selective medium (SC-ALW) containing varying concentrations of ligand. Yeast growth would indicate ligand binding and subsequent hVDR activation.

hVDR was tested with $1,25(OH)_2D_3$, lithocholic acid (LCA), and cholecalciferol. Cholecalciferol is an intermediate in the biosynthesis of $1,25(OH)_2D_3$ and does not activate wild-type hVDR, therefore this ligand is used as a negative control [19]. The ligand concentration ranges tested were between 0.001 µM to 1 µM for $1,25(OH)_2D_3$ and 0.01 µM to 10 µM for LCA and cholecalciferol.

EC_{50} values, the ligand concentration at which 50% of the maximal response is observed, are important for determining proper concentrations of a drug needed to be administered to achieve the desired effect [20]. Gal4 is as the positive control because this protein is a ligand activated transcription factor and should therefore grow in the presence or absence of ligand. As expected, hVDR displays an EC_{50} value of 5 nM $1,25(OH)_2D_3$ in adenine selective media (SC-ALW) with an 8-fold activation based on growth, which is comparable to literature values in mammalian cells with an EC_{50} value

of 1 nM (Figure 2.3A) [21]. No growth was observed with LCA and cholecalciferol in adenine selective media (Figure 2.3B); therefore, hVDR was then tested in histidine selective media with LCA and cholecalciferol to determine whether a less stringent selection marker would allow for growth with these ligands.

The *HIS3* gene is considered a "leaky gene" in PJ69-4A, due to residual basal growth observed in the absence of ligand. Basal growth is reduced by adding 3-amino 1,2,4 triazole (3-AT) to the media, which inhibits the enzyme encoded by the *HIS3* gene, imidazoleglycerol-phosphate dehydratase [22]. A liquid quantitation assay was performed to determine the concentration of 3-AT at which basal activity would be reduced, allowing for ligand-activated growth. Based on these results, 0.1 mM 3-AT was determined to be the optimal concentration. At 0.1 mM 3-AT, ligand-activated growth was observed as well as no basal growth being observed [23].

As shown in Figure 2.3C, LCA has an EC_{50} of > 5μM for wild-type hVDR in histidine selective media (SC-HLW and 0.1 mM 3AT), with 3-fold activation compared to the absence of ligand. As expected, ligand-activated growth was not observed with cholecalciferol at any ligand concentration (Figure 2.3A and C). Gal4 was used as the ligand-independent control, and colonies containing Gal4 grew in the absence and presence of ligand. These results validate the application of hVDR in chemical complementation and are similar to literature values in mammalian cells, where LCA shows activation at 100 μM and cholecalciferol shows no activation [24]. Thus, chemical complementation is a valuable tool for evaluating the hVDR protein libraries that were created.

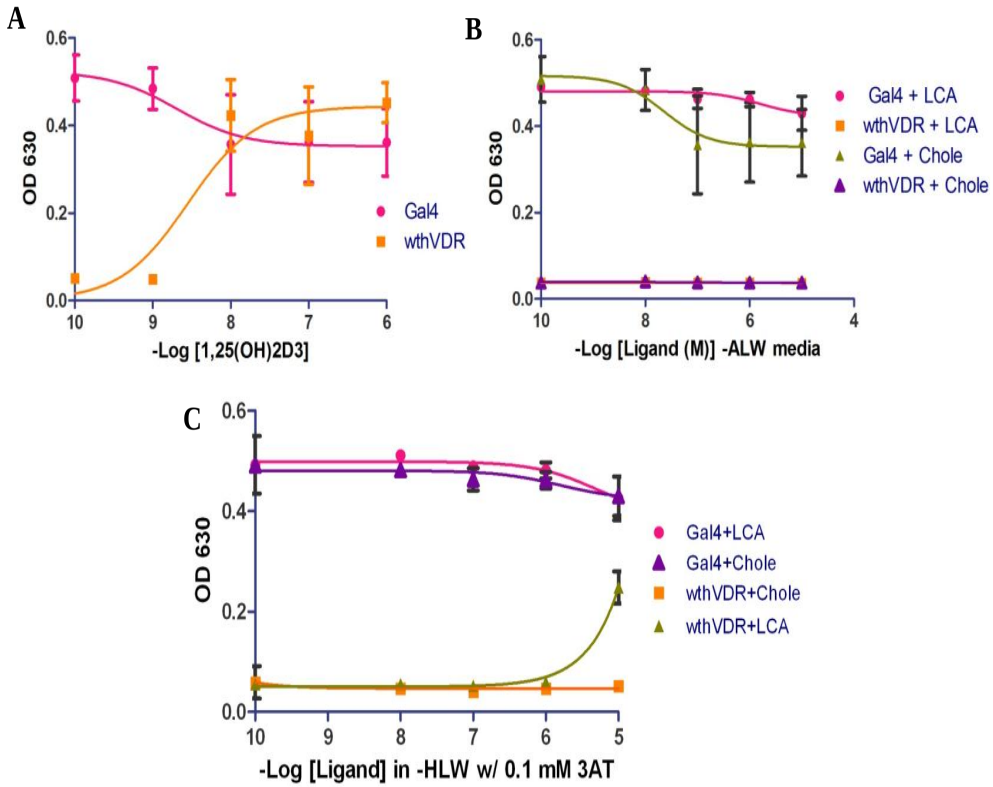

Figure 2.3: Testing wild-type hVDR with various ligands. Gal4 displays liagnd-independent growth will all of the ligands. (A) Wild-type hVDR displays ligand-activated growth at 10 nM $1,25(OH)_2D_3$ in adenine selective media. (B) No ligand-activated growth is observed with lithocholic acid and cholecalciferol in adenine selective media. (C) Wild-type hVDR displays ligand-activated growth at 10 µM lithocholic acid and no growth is observed with cholecalciferol in histidine selective media with 0.1 mM 3-AT. The y-axis represents yeast growth, while the x-axis represents increasing ligand concentration.

2.2 hVDR Library 1

2.2.1 Background: Library Design

A library was designed based on the crystal structure of hVDR bound to $1,25(OH)_2D_3$ to determine the mutational tolerance of the residues within the ligand binding pocket. The crystal structures were analyzed to determine residues in the LBP that were in direct contact with the ligand and within four angstroms [3]. Four angstroms was chosen as the maximal distance for contacts between the ligand and receptor in order to identify residues that are within the reach of a hydrogen bond (~ 1.5 Å- 3.5 Å), an electrostatic interaction, and a van der Waals interaction with the ligand. Twenty amino acid residues were found within this distance range. Further analysis included a consideration of previous mutational analysis as well as sequence alignments of the human VDR and other VDR orthologs to determine suitable residues for the library generation (Table 2.1). The partial sequence alignment, as shown in Table 2.1, shows that most hVDR LBD residues are conserved. Therefore, the residues chosen to be mutated were based on the crystal structure of $1,25(OH)_2D_3$ with wild-type hVDR and previous mutational analysis. Residues were chosen for mutation based on the structural analysis as well as residues that were known to form important contacts with $1,25(OH)_2D_3$ in an effort to determine if what mutations would allow for wild-type activity with $1,25(OH)_2D_3$. Seven of the twenty residues were selected for mutagenesis in Library 1 in order to minimize the library size (Table 2.2). The seven residues were L227, L230, L233, V234, S237, H305, and H397.

2.2.2 hVDR Library 1 Design

Table 2.1: Sample alignment of VDR orthologs.

	379								387										397	
VDR_Human	L	Y	A	K	M	I	Q	K	L	A	D	L	R	S	L	N	E	E	H	S
VDR_Chicken	L	Y	A	K	M	I	Q	K	L	A	D	L	R	S	L	N	E	E	H	S
VDR_Mouse	L	Y	A	K	M	I	Q	K	L	A	D	L	R	S	L	N	E	E	H	S
VDR_Rat	L	Y	A	K	M	I	Q	K	L	A	D	L	R	S	L	N	E	E	H	S
VDR_Frog	L	Y	A	K	M	I	Q	K	L	A	D	L	R	S	L	N	E	E	H	S
VDR_Zebrafish	L	Y	A	K	M	I	Q	K	L	A	D	L	R	S	L	N	E	E	H	S
VDR_Halibut	L	Y	A	K	M	I	Q	K	L	A	D	L	R	S	L	N	E	E	H	S
VDR_Lamprey	L	Y	A	K	M	I	Q	K	L	A	D	L	R	S	L	N	E	E	H	S
VDR_Quail	L	Y	A	K	M	I	Q	K	L	A	D	L	R	S	L	N	E	E	H	S
VDR_Tamarin	L	Y	A	K	M	I	Q	K	L	A	D	L	R	S	L	N	E	E	H	S

	399									408									418
VDR_Human	K	Q	Y	R	C	L	S	F	Q	P	E	C	S	M	K	L	T	P	L	V
VDR_Chicken	K	Q	Y	R	C	L	S	F	Q	P	E	H	S	M	Q	L	T	P	L	V
VDR_Mouse	K	Q	Y	R	S	L	S	F	Q	P	E	N	S	M	K	L	T	P	L	V
VDR_Rat	K	Q	Y	R	S	L	S	F	Q	P	E	N	S	M	K	L	T	P	L	V
VDR_Frog	K	Q	Y	R	S	I	S	F	L	P	E	H	S	M	K	L	T	P	L	M
VDR_Zebrafish	K	Q	Y	R	S	L	S	F	Q	P	E	H	S	M	Q	L	T	P	L	V
VDR_Halibut	K	Q	Y	R	S	L	S	F	R	P	E	H	S	M	Q	L	T	P	L	V
VDR_Lamprey	K	Q	Y	R	S	L	S	F	Q	P	E	H	S	M	Q	L	T	P	L	V
VDR_Quail	K	Q	Y	R	C	L	S	F	Q	P	E	H	S	M	Q	L	T	P	L	V
VDR_Tamarin	K	Q	Y	R	C	L	S	F	Q	P	E	S	S	M	K	L	T	P	L	V

The red residues are examples of conservation among different VDR orthologs. The orthologs are listed in the far left column. Most residues are conserved between the different orthologs.

Table 2.2: Library 1 Design. Residues and chosen mutations are listed. The types of changes caused by the mutations are also listed.

Amino Acid	Possible Changes	Goal of Mutations
L227	LIV_2FM	Hydrophobic/ volume
L230	LIV_2FM	Hydrophobic/ volume
L233	LIV_2FM	Hydrophobic/ volume
V234	LIV_2FM	Hydrophobic/ volume
S237	LIV_2FMAST	Hydrophobic/ volume/ polar
H305	Saturated	All
H397	HYNKQstop	Polar/ charged

Library 1 was constructed to determine whether changing the volume of the pocket would modify the binding properties of the receptor. In addition to residues with volume changes, the chemical properties of the residues were also modified. For example, if important hydrogen bonding residues are changed to non-polar residues, an effect in ligand sensitivity and preference should be observed, due to the loss of hydrogen bonding with $1,25(OH)_2D_3$. L227, L230, L233, and V234 were targeted for mutation based on their proximity to the ligand and were modified to other hydrophobic residues, such as leucine, isoleucine, valine, phenylalanine, and methionine, to alter the pocket volume (Table 2.2). S237, which forms a hydrogen bond with the 1α-hydroxyl group of $1,25(OH)_2D_3$, was modified to threonine and the more hydrophobic residues leucine, isoleucine, valine, phenylalanine, methionine, alanine [3]. Additionally, two histidine residues were also targeted: H305, which is located in a loop between helices 6 and 7, and H397, which is located in helix 11. Previously, the role of H305 in the LBD of hVDR was determined to be significant for the binding of $1,25(OH)_2D_3$, as H305 is thought to form a hydrogen bond with the 25-hydroxyl group of the ligand. However, Mizwicki *et al.* showed that an H305F mutant had no affect on activation by $1,25(OH)_2D_3$ [21, 25]. This residue was saturated to determine the significance of this position [3, 5]. In the case of H397, which forms a crucial hydrogen bond with the 25-hydroxyl group of $1,25(OH)_2D_3$, tyrosine, lysine, glutamine, and asparagine were chosen to determine whether a charged or polar residue would be tolerated at this position (Table 2.1) [3, 26]. The combination of these designed mutations created a theoretical library size of 5.0×10^5 variants.

2.2.3 Experimental Techniques for Designing Library

In order to create and analyze the libraries in yeast, three components are needed: an insert cassette containing hVDR variants, a background vector that contains sequences of homology with the insert cassette, and the coactivator/GAD fusion protein (pGAD10BAACTR). Protein variants are created *in vivo* via homologous recombination, with an insert cassette and background vector with complimentary ends that allow for homologous recombination. The insert cassette contains the selected mutations and the background vector is used for recombination and contains the Gal4 DBD.

The background vector contains a random DNA sequence that encodes several stop codons to eliminate the wild-type hVDR background when transformed into yeast. The background plasmid was created by initially inserting a *Sac*II site via site-directed mutagenesis at the beginning of the hVDR LBD gene in pGBDhVDR and a *Kpn*I site after the stop codon, creating pGBDhVDRSacIIKpnI. pGBDhVDRSacIIKpnI was digested with *Sac*II and *Kpn*I, removing an 850 bp piece of the hVDR LBD. DNA from a separate plasmid, pMSCVeGFP, was amplified with *Sac*II and *Kpn*I restriction sites at the ends, and then ligated in between the *Sac*II and *Kpn*I sites, creating the background vector (pGBDhVDRbackground) with stop codons in the reading frame (Figure 2.4). The background vector is linearized, allowing for recombination with the insert cassette, which contains complementary ends to the background vector.

Individual oligonucleotide primers (oligos) containing degenerate base codes at the specific codon locations were designed and combined through hybridization and overlap extension PCR, to generate a library of hVDR variant insert cassettes [27]. Oligonucleotides containing the randomized codons for the amino acid residues were ordered and designed, such that oligo 1 and oligo 8 contain the complementary ends to

the background vector. As illustrated in Figure 2.5, the complimentary ends of oligos 1-4 and 5-8 were created via hybridization and were amplified via PCR to create cassettes A and B, which were then allowed to hybridize and amplify to create the complete insert cassette. The resulting insert cassette was 1014 bp.

The plasmid containing a fusion of the human co-activator ACTR and the Gal4 activation domain with a leucine marker (pGAD10BAACTR) as well as the background vector and the insert cassette was transformed into yeast. The yeast tranformants were plated onto adenine selective plates (SC-ALW) with LCA and cholecalciferol [14]. Yeast growth will be observed in the selective media if the ligand is able to bind and activate the hVDR variant. Transformants were also plated onto synthetic complete media lacking leucine and tryptophan (SC-LW), to determine transformation efficiency and ensure that recombination has occurred. The transformation efficiency was determined to be 1.0×10^6 colony forming units/µg of vector cassette.

2.2.4 Results of hVDR Library 1

A library size of 1.0×10^5 variants was observed when plated onto non- selective plates (-LW) and selective plates with LCA (SC-ALW). Two colonies grew on the adenine selective plates (SC-ALW), which were then streaked onto selective plates without ligand in order to confirm that growth was due to ligand activation, and not a constitutively active hVDR variant (Figure 2.6). In yeast, constitutive activity could result from the NR binding an endogenous ligand/metabolite, leading to an active receptor and turning on transcription of the reporter gene. Constitutive activity could also be due to one or more of the introduced mutations causing an active confirmation of the

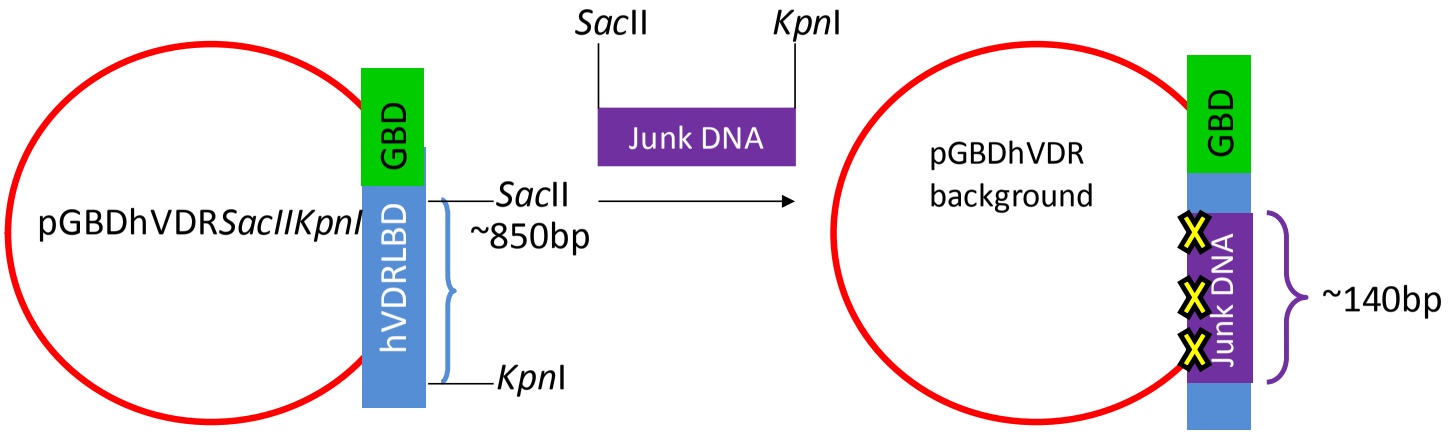

Figure 2.4: Creation of hVDR background plasmid. A section of the hVDRLBD is removed and replaced with a piece of random DNA that encodes for several stop codons. The background plasmid will eliminate wild-type contamination in the chemical complementation system. (The yellow X's represent stop codons.)

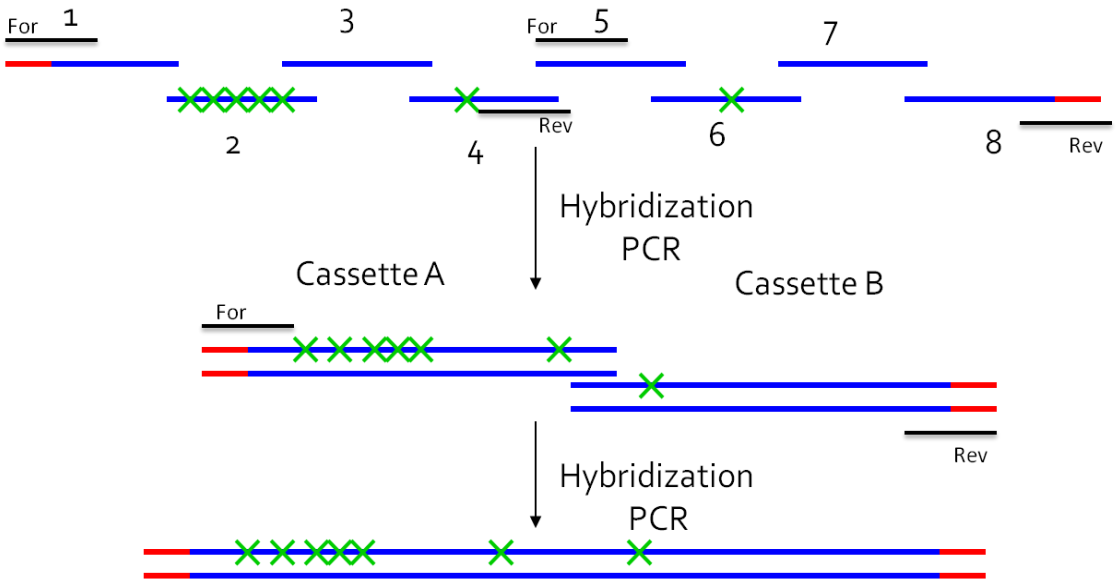

Figure 2.5 Using overlap extension PCR to create insert cassettes. The individual oligos contain complimentary ends with the next oligo. Therefore, the oligos can be pieced together using hybridization and PCR. Oligos 1-4 and 5-8 were hybridized and amplified to create cassettes A and B. Cassettes A and B were then hybridized and amplified to create the full insert cassette. Forward (For) and reverse (Rev) primers were used for the amplification steps.

Adenine selective plate
No Ligand

Adenine selective plate
10^{-5} M Lithocholic Acid

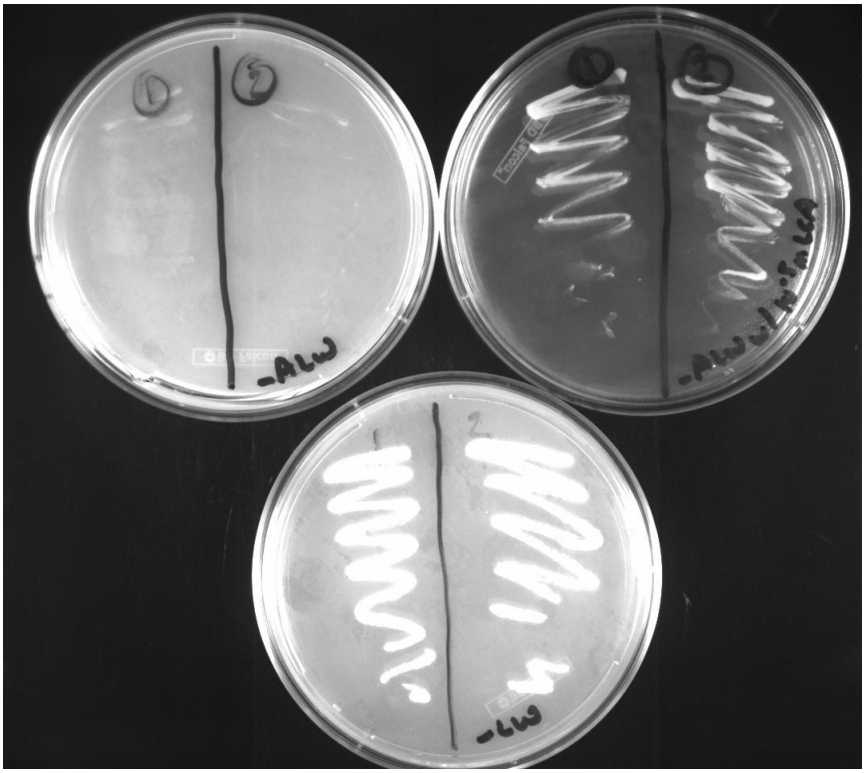

Figure 2.6: Streaking of Library #1 colonies. Selective colonies were streaked onto adenine selective plates (-ALW) without ligand to test for constitutive activity. Constitutive activity is growth in the absence of ligand. No growth was observed on the ligand plate without ligand. The colonies show ligand-activated growth with 10^{-5} M LCA in adenine selective media.

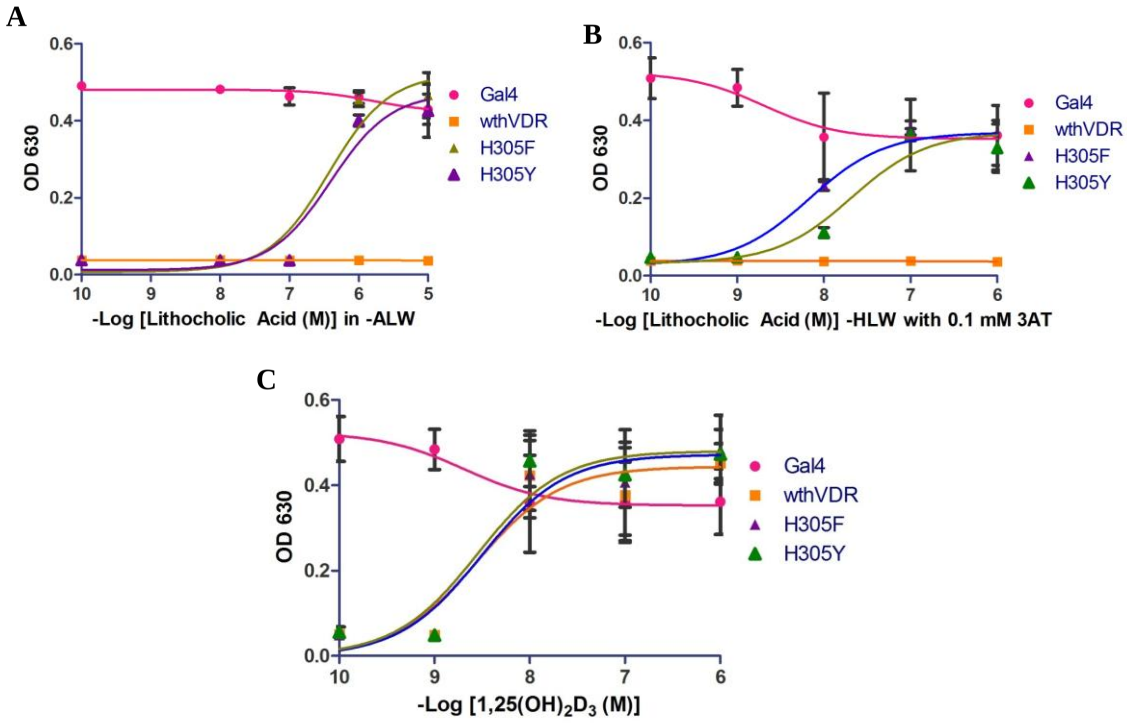

Figure 2.7: Testing hVDR variants in chemical complementation. (A) Ligand-activated growth is observed for the two variants at 1 µM LCA in comparison to no growth for wild-type hVDR in adenine selective media. (B) The two variants display liagnd-activted growth at 100 nM LCA in histidine selective media with 0.1 mM 3-AT. (C) Wild-type hVDR and variants display ligand-activated growth at 10 nM $1,25(OH)_2D_3$ in adenine selective media.

receptor in the absence of ligand. Streaking results indicated that the two colonies were not constitutively active, prompting further testing in a liquid quantitation assay using chemical complementation as mentioned previously.

When tested with LCA and $1,25(OH)_2D_3$, both variants were found to have enhanced sensitivity to LCA when compared to wild-type hVDR. As shown in Figure 2.7A, these two variants display an EC_{50} value in adenine selective media (-ALW) of 0.5 µM LCA with 12-fold activation, while wild-type hVDR showed no growth in adenine selective media. Both variants display an EC_{50} value of 0.5 µM LCA in histidine selective media with a 9-fold activation compared to wild-type hVDR, which displays an EC_{50} value of >5 µM LCA with a 3-fold activation in histidine selective media, and a 10-fold increase in sensitivity (Figure 2.7B). Wild-type hVDR and the variants were analyzed for activation with $1,25(OH)_2D_3$, for which EC_{50} values of 5 nM and 11 fold-activation were observed in adenine selective media for each (Figure 2.7C).

Variants were rescued from the non-selective and selective plates and sequenced. The two selective variants both contained a mutation at H305, one to phenylalanine and the other to tyrosine. Sequencing results indicated that of the 18 non-selective variants, 67% contain mutations at residues H305 and H397. 94% of the variants contained insertions, deletions, and non-designed mutations, leading to inactive variants (Table 2.3). Based on the results, increased bulk of the side chains, specifically histidine to phenylalanine or tyrosine, leads to ligand activation, and changing the polarity at H305 does not seem to decrease ligand activation, confirming the results observed by Mizwicki *et al.* Mizwicki *et al.* showed that mutating H305 to a phenylalanine does not affect

Table 2.3: Library #1 Sequencing Results. Residues that were not changed are shown as wt. Mutations are shown along with the codon change. As can be observed, the majority of the variants maintained the wt residue at residues L227-S237. Diversity was only observed at H305 and H397.

	LIVFM	**LIVFM**	**LIVFM**	**LIVFM**	**LIVFMAST**	**Saturated**	**HYNKQstop**
	L227	**L230**	**L233**	**V234**	**S237**	**H305**	**H397**
Variant				**Non-selective**			
2*	xTC	V(GTG)	V(GTG)	F(TTC)	wt	D(GAC)	H(CAT)
3*	wt	wt	wt	wt	wt	R(AGG)	N(AAT)
4*	F(TTC)	I(ATC)	V(GTC)	M(ATG)	S(TCC)	C(TGT)	Y(TAT)
5*	wt	wt	wt	wt	wt	I(ATT)	H(CAT)
7	wt	wt	wt	wt	wt	V(GTT)	wt
8*	wt	wt	wt	wt	wt	wt	H(CAT)
9*	wt	wt	wt	wt	wt	R(CGC)	Y(TAT)
10*	wt	wt	wt	wt	wt	R(AGG)	H(CAT)
12*	wt	wt	wt	wt	wt	V(GTT)	N(AAT)
13*	wt	wt	wt	wt	wt	I(ATT)	Q(CAG)
				Selective			
Variant							
1	wt	wt	wt	wt	wt	Y(TAT)	wt
2	wt	wt	wt	wt	wt	F(TTC)	wt

*= Insertions and/or deletions in variants. Blue= positions where diversity was observed

activation with $1,25(OH)_2D_3$, implying that a hydrogen bond is not essential at H305 for activation [21, 25]. Although ligand-activated variants were not discovered with mutations at the other residues chosen in the library, a conclusion of tolerance at these positions could not be made due to the fact that a complete library of variants was not achieved.

LCA was docked into wild-type hVDR and H305F using AutoDock Vina in order to visualize the influence of these mutations on ligand binding through *in silico* methods (Figure 2.8) [28]. The structures of the hVDR variants were prepared *in silico* using the program TRITON 4.0.0 (National Centre for Biomolecular Research, Czech Republic) and its external program MODELLER (National Centre for Biomolecular Research, Czech Republic). The computational site-directed mutagenesis method, which is based on using the wild-type protein for homology modeling was implemented [29, 30]. Homology modeling analyzes all of the crystal structures in the PDB, creating an algorithm of the residue conformations based upon residues in proximity to the residue of interest. For example, based on all of the crystal structures, leucine has been found to prefer a certain conformation when this residue is next to an alanine or a tryptophan.

To develop models for analysis, the wild-type hVDR crystal structure (PDB:1DB1) was used as the template [3]. Variants were prepared for docking using the UCSF CHIMERA- interactive molecular graphics program by: (1) removing the ligand and water molecules, (2) adding polar hydrogens, and (3) assigning Gasteiger charges, which sets the partial charge property of each atom [31]. Ligands were created using ChemBioDraw Ultra 11.0 and ChemBio3D Ultra 11.0 (Cambridge Soft, USA) and modified with the AutoDockTools by adding Gasteiger charges, setting the partial charge

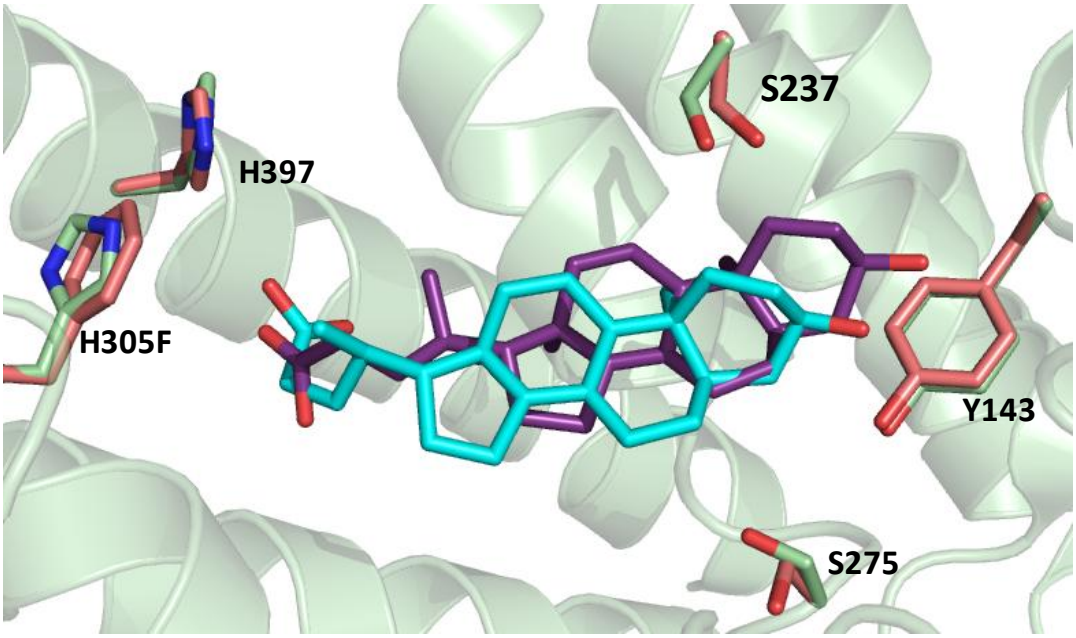

Figure 2.8: wthVDR and H305F docked with lithocholic acid. Wild-type hVDR= green (residues) and cyan (ligand). H305F= salmon (residues) and purple (ligand). A shift in the carboxyl group of LCA can be seen in the H305F docking. Also, a shift of the ligand towards Y143 can be seen in the H305F docking when compared to LCA docked to wild-type hVDR.

property of each atom [32]. AutoDock Vina was used to perform docking simulations with default parameters, such as search area and time spent searching the area for conformations [28]. The solutions exhibiting the lowest free energy of binding were analyzed.

In silico modeling of $1,25(OH)_2D_3$ was used as the control to verify that the modeling method works. *In silico* modeling results indicated that the H305F and H305Y variants cause a shift in the position of LCA in the ligand binding pocket. When comparing the superimposition of the docked structures of wild-type hVDR and H305F bound to lithocholic acid (Figure 2.8), a shift in the LCA ring system towards Y143 and a repositioning of the carboxyl group of LCA is observed. The bulkier residues, phenylalanine (189.9 $Å^3$) or tyrosine (196.3 $Å^3$), in the place of histidine (153.2 $Å^3$), lead to an increase in volume at position 305 [33]. This increase seems to contribute to the repositioning and shifting of the ligand, resulting in closer interactions between residues in the pocket and LCA, thus enhancing sensitivity with this ligand [25, 33]. H305F and H305Y were subjected to further mutagenesis to engineer a hVDR variant capable of binding and activation in response to cholecalciferol.

Based upon previous work by Mizwicki *et al.* the replacement of H305 with a non-polar aromatic residue, phenylalanine, was expected to affect sensitivity to LCA, as their observations indicate that the H305F mutation does not affect activation with $1,25(OH)_2D_3$ [21, 25]. H305F has also been shown to reverse ligand activity by turning the hVDR antagonist, (23*S*)-25-dehydro-1-(OH)-vitamin D3-26,23-lactone (MK), into a superagonist, which means activation with MK occurs at a lower concentration relative to wild-type hVDR activation with $1,25(OH)_2D_3$. This is believed to occur because MK

requires less space in the LBP than $1,25(OH)_2D_3$ and therefore does not create the necessary contacts with H305 that are needed to obtain activation [25]. However, increasing the bulk at this position with a phenylalanine allows for additional contacts with the 305 residue, due to the decrease of LBP volume, resulting in ligand activation. The same concept is believed to apply with LCA, as the smaller ligand LCA is predicted to have fewer contacts with H305 but increased contacts with H305F. The H305F variant of hVDR displays an increased sensitivity toward LCA in comparison to wild-type. As LCA is smaller than $1,25(OH)_2D_3$, the appearance is that bulkier residues, such as phenylalanine or tyrosine may be optimizing the interactions with LCA in comparison to H305. These mutations do not affect activation with $1,25(OH)_2D_3$, suggesting that the ligand/receptor contacts are maintained for binding and activation.

2.3 hVDR Library 2

As mentioned previously, twenty residues were found within four angstroms of $1,25(OH)_2D_3$ in the hVDR ligand binding pocket. However, Library # 1 contained only seven residues in order to minimize library size. Accordingly, five more residues from the twenty were chosen as targets in library 2: Y143, R274, S275, S278, and W286 (Figure 2.1, Table 2.4). The goal of this library was to challenge the volume, shape, and polarity of the pocket with residues that could not be included in Library 1. All of the residues with the exception of W286 are believed to be involved in hydrogen bonding with the ligand based upon previous alanine scanning mutagenesis. The mutations chosen for each residue can be found in Table 2.5. Y143, which forms a hydrogen bond to the 3-hydroxyl group of $1,25(OH)_2D_3$, was subjected to saturation mutagenesis to test all chemical and physical properties of the different residues. R274, S275, and S278,

Table 2.4: Library 2 Design. Residues and chosen mutations are listed. The types of changes caused by the mutations are also listed.

Amino Acid	Mutations	Goal of Mutations
Y143	saturated	all
R274	$LIV_2FMA_2S_2T_2CRWG_2$	Hydrophobic/volume/polar
S275	$LIV_2FMA_2S_2T_2$	Hydrophobic/volume/polar
S278	$LIV_2FMA_2S_2T_2$	Hydrophobic/volume/polar
W286	F_2WLC_2	Hydrophobic/volume

which also form hydrogen bonds to $1,25(OH)_2D_3$, were mutated to hydrophobic residues in order to modify the polarity of the pocket. W286 was mutated to two non-aromatic residues (leucine and cysteine) and two aromatic residues (phenylalanine and tryptophan). The theoretical library size (based on amino acids) was 6.1×10^4 variants.

The same approach was used to create these cassettes as was used with Library 1, and these variants were transformed into yeast. A library size of 1.1×10^3 variants was observed. However, no variants displayed growth on histidine selective plates. To determine the library's diversity, six non-selective colonies were sequenced. Sequencing analysis revealed that diversity was observed, and approximately 66% of the variants displayed insertions, deletions, and undesigned mutations (Table 2.5), indicating a flaw in the overlap extension and hybridization protocol for creating the insert cassettes.

2.4 hVDR Library 3

Library 3 focused primarily on the role of hydrogen bonding in ligand activation of hVDR (Figure 2.1). All hydrogen bonding residues from Libraries 1 and 2 were combined into one library to assess the effect when these residues were mutated to non-hydrogen bonding residues. The designed mutations remained the same as in the previous libraries: Y143 and H305 were subjected to saturation mutagenesis, and S237, R274, and S278 were mutated to hydrophobic residues. H397 was mutated to tyrosine, lysine, glutamine, and asparagine (Tables 2.2, 2.4). The theoretical library size was 2.0×10^6 variants and a library size of 2.3×10^3 variants was observed, after transforming the variants into yeast.

Two variants were observed on histidine selective plates with LCA. These variants were streaked onto histidine selective plates (SC-HLW with 0.1 mM 3AT) without

Table 2.5: Library #2 Sequencing. Residues that were not changed are shown as wt. Mutations are shown along with the codon change. As can be observed, diversity is observed at the positions that were chosen for mutation.

	Saturated	LIVFMASTCRWG	LIVFMAST	LIVFMAST	FWLC
Variant	**Y143**	**R274**	**S275**	**S278**	**W286**
1	stop	M(ATG)	wt	L(TTG)	C(TGT)
2	S(TCG)	A(GCC)	V(GTC)	F(TTT)	L(TTG)
3	E(GAG)	V(GTC)	A(GCT)	T(ACC)	L(TTG)
4*	T(ACG)	T(ACC)	L(CTG)	A(GCG)	wt
5	P(CCG)	W(TGG)	I(ATC)	V(GTG)	C(TGC)
6	A(GCT)	M(ATG)	V(GTG)	I(ATC)	wt

*= Insertions and/or deletions found in variant. Blue = diversity seen at these positions.

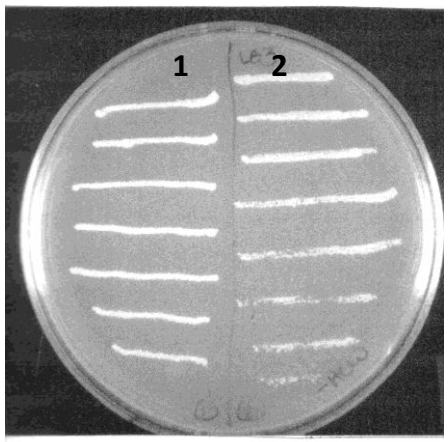

Figure 2.9: Library # 3 Streaking. Streaking selective colonies onto histidine selective media without ligand (-HLW w/ 0.1 mM 3-AT) to test for constitutive activity. Both colonies displayed growth and are therefore constitutively active.

ligand. As seen in Figure 2.9, both colonies were constitutively active showing growth in the absence of ligand. Based upon sequencing of variants from the library, 84% of the sequenced colonies contained insertions, deletions, or undesigned mutations (Table 2.6). Both constitutively active variants were terminated prior to the stop codon at Q239, leading to a truncated protein. This was surprising, considering that helix 12 is known to be crucial for NR receptor activation.

2.5 Summary

hVDR was engineered for a better understanding of the mutational tolerance of the pocket, as well as further characterizing the residues necessary for ligand activation. Three libraries were created in an attempt to better understand the interactions between hVDR and LCA. In Library 1, two variants, H305F and H305Y, were discovered with an increased sensitivity towards LCA, and Libraries 2 and 3 did not produce any variants capable of ligand activation. However, diversity in the libraries was observed, along with insertions, deletions, and undesigned mutations. Therefore, chemical complementation was shown to be a useful tool to analyze hVDR libraries. The two variants, H305F and H305Y, give some insight into the tolerance of position H305 in terms of ligand binding and activation. In an attempt to engineer hVDR variants capable of binding a novel small molecule, the two variants from Library 1 were subjected to multiple cycles of mutagenesis for protein engineering.

2.6 Methods and Materials

2.6.1 Ligands

Table 2.6: Library #3 Sequencing.

	Saturated	LIVFMAST	LIVFMASTCRWG	LIVFMAST	Saturated	HYNKQstop
	Y143	**S237**	**R274**	**S278**	**H305**	**H397**
Variant			**Non-selective**			
1*	T(ACG)	wt	T(ACC)	A(GCG)	wt	wt
2	A(GCT)	wt	M(ATG)	I(ATC)	wt	wt
3*	F(TTT)	A(GCG)	G(GGC)	V(GTC)	G(GGT)	Q(CAG)
4*	G(GGG)	M(ATG)	wt	wt	Y(TAT)	K(AAG)
5*	wt	F(TTC)	G(GGC)	I(ATC)	Y(TAT)	Q(CAG)
Variant			**Selective**			
1**	wt	A(GCC)	S(AGC)	L(TTG)	T(ACT)	N(AAT)
2**	A(GCG)	I(ATC)	M(ATG)	A(GCC)	stop	N(AAT)

*= Insertions and/or deletions in variant. **= Q239stop. Residues shown in blue are not found in the non-selective sequencing. All of the positions show some diversity, but not complete diversity.

Lithocholic acid, cholecalciferol, and 1α,25-dihydroxyvitamin D_3 were ordered from MP Biomedicals, LLC (Solon, OH), Sigma-Aldrich (St. Louis, MO), and BIOMOL (Plymouth Meeting, PA), respectively. 10 mM stocks of lithocholic acid and cholecalciferol and a 13.3 μM stock of 1α,25-dihydroxyvitamin D_3 were made with 80% ethanol: 20% DMSO as the solvent and stored at 4 °C.

2.6.2 Creating Designed hVDR Library

To create the hVDR insert cassette variants, oligonucleotides with overlapping ends, containing degenerate base codes at the chosen positions, were pieced together using overlap extension PCR (Figure 2.4, Tables 2.7, 2.8, 2.9) [27]. 100 ng of each oligo and 125 ng of each primer were combined in a PCR tube and subjected to a PCR program with the following conditions: 95 °C for 1 min, 59 °C for 1 min, 72 °C for 2 min, 15 cycles, 95 °C for 1 min, 56.6 °C for 1 min, 72 °C for 2 min, 15 cycles. The resulting insert cassette was 1014 bp.

To create a background vector cassette, site-directed mutagenesis was used to insert *Sac*II and *Kpn*I sites into pGBDhVDR 810 bp apart, resulting in the pGBDhVDRSacIIKpnI plasmid. This plasmid was digested with the two restriction enzymes, removing 810 bp of the wild-type VDR gene from the plasmid. A 145 bp segment of random DNA from another plasmid, pMSCVeGFP, was digested and ligated between the *Sac*II and *Kpn*I sites (pGBDhVDRSacIIKpnI), generating 3 stop codons in the open reading frame. The resulting plasmid was named pGBDhVDRbackground and confirmed by sequencing (Operon, USA).

2.6.3 Yeast Transformation

Table 2.7: Library 1 oligonucleotides with degenerate base codes at mutation sites.	
Oligo #	**Oligo Sequence**
Oligo1For	gatcctgaagcggaaggagg
Oligo1Rev	atggagagctgggacagctctag
Oligo2Rev	agcaaagccaatgacctttggatSRHgtaactSAHSAHgtcagcSAHgtggggSAHcatggagagctgggacagctctag
Oligo3For	ccaaaaggtcattggctttgct
Oligo3 Rev	ggctttggtcacgtcactgacg
Oligo4Rev	tcccacctggaacttgatgaggggctcaatcagctccaggctVNNtccggctttggtcacgtcactgacg
Oligo5For	cctcatcaagttccaggtggga
Oligo5Rev	cattgaggctgcgcaggtc
Oligo6Rev	caggctggaaggagagaggcagcggtactgcttggaMTDctcctcattgaggctgcgcaggtc
Oligo7For	ctgcctctccttccagcctgagtgcagcatgaagctaacgccccttgtgctcgaagtgtttggcaatgagat
Oligo8Rev	gtccaggcagggtggccagaacgggtgggcacaaaggatggactagttcaggagatctcattgccaaacacttcgag
Capital letters represent degenerate base codes at the chosen residues. The primers that are around twenty base pairs are used to amplify parts of the VDRLBD that do not contain mutations.	

Table 2.8: Library 2 Oligonucleotides with degenerate base codes at mutation sites	
Oligo #	**Oligo Sequence**
Oligo1For	gatcctgaagcggaaggagg
Oligo1Rev	cgtccagcagtatggcaatga
Oligo2Rev	ctggcagaagtcggagtaggtggggtcMNNggtcttatggtgggcgtccagcagtatggcaatga
Oligo3For	cacctactccgacttctgccag
Oligo3 Rev	caacatgatgacctcaatggcac
Oligo4Rev	ttgtagtcttggttgccacaggtVMAggacatgtcgtccatggtgaaSRHctcattSRHSVHcaacatgatgacctcaatggcac
Oligo5For	acctgtggcaaccaagactacaa
Oligo5Rev	atctcattgccaaacacttcgag
Oligo6Rev	gtccaggcagggtggccagaacgggtgggcacaaaggatggactagttcaggagatctcattgccaaacacttcgag
Capital letters represent degenerate base codes at the chosen residues. The primers that are around twenty base pairs are used to amplify parts of the VDRLBD that do not contain mutations.	

Table 2.9: Library #3 Oligonucleotides with Degenerate base codes at positions of mutation.	
Oligo #	**Oligo Sequence**
Oligo1For	gatcctgaagcggaaggagg
Oligo1Rev	cgtccagcagtatggcaatga
Oligo2Rev	ctggcagaagtcggagtaggtggggtcMNNggtcttatggtgggcgtccagcagtatggcaatga
Oligo3For	ctgaccaggtcagccaggtg
Oligo3 Rev	cacctactccgacttctgccag
Oligo4Rev	aggtgaggtctctgaatcctggtatcatcttagcaaagccaatgacctttggatSRHgtaactgaccaggtcagccaggtg
Oligo5For	ataccaggattcagagacctcacctctgaggaccagatcgtactgctgaagtcaagtgccattgaggtcatcatgttg
Oligo6Rev	ttgtagtcttggttgccacaggtccaggacatgtcgtccatggtgaaSRHctcattggaSVHcaacatgatgacctcaatggcac
Oligo7For	acctgtggcaaccaagactacaagtaccgcgtcagtgacgtgaccaaagccggaNNKagcctggagctgattgagcc
Oligo8For	agcctggagctgattgagcc
Oligo8Rev	cattgaggctgcgcaggtc
Oligo9For	gacctgcgcagcctcaatgaggagHAKtccaagcagtaccgctgcctctccttccagcctg
Oligo10For	tgcctctccttccagcctg
Oligo10Rev	gtccaggcagggtggcc
Capital letters represent degenerate base codes at the chosen residues. The primers that are around twenty base pairs are used to amplify parts of the VDRLBD that do not contain mutations.	

Using the 1x TRAFCO yeast transformation protocol [34], 0.1 µg of the digested background plasmid (pGBDhVDRbackground) and 0.9 µg of insert cassette were transformed into the yeast strain PJ69-4A. A plasmid containing a fusion of the human co-activator ACTR and the Gal4 activation domain with a leucine marker was also co-transformed [14]. Transformants were plated onto synthetic complete (SC) agar plates lacking adenine, leucine, and tryptophan (SC-ALW) with various ligands. Synthetic complete plates are made using yeast nitrogenous base, ammonium sulfate, dextrose anhydrous, and dropout powder. The dropout powder used does not contain uracil, tryptophan, adenine, leucine, or histidine. Transformants were also plated without ligand onto synthetic complete agar plates lacking leucine and tryptophan (SC-LW) to determine transformation efficiency. Transformation efficiency was calculated based upon the number of colonies on the plate divided by the µg of vector cassette that was plated onto each plate.

2.6.4 Liquid Quantitation Assay of hVDR in Yeast

Variants were tested in liquid quantitation assays in 96-well plates with media lacking adenine, leucine, and tryptophan (SC-ALW), with or without LCA, cholecalciferol, or $1,25(OH)_2D_3$ at varying concentrations (ranging from 0.01 µM - 10 µM for LCA and cholecalciferol, 1 nM - 1 µM for $1,25(OH)_2D_3$). A 4:1 ratio of media (SC-ALW): cells (yeast resuspended in water) were aliquoted into 96-well plates. Plates were incubated at 30°C, with shaking at 170 rpm. Optical density (OD) readings at 630 nm were recorded at 0, 24, and 48 hours to measure turbidity.

2.7 References

1. Malloy PJ, Eccleshall TR, Gross C, VanMaldergem L, Bouillon R, Feldman D: **Hereditary vitamin D resistant rickets caused by a novel mutation in the**

vitamin D receptor that results in decreased affinity for hormone and cellular hyporesponsiveness. *J Clin Invest* 1997, **99**(2):297-304.

2. Rochel N, Hourai S, Moras D: **Crystal structure of hereditary vitamin D-resistant rickets--associated mutant H305Q of vitamin D nuclear receptor bound to its natural ligand**. *J Steroid Biochem Mol Biol*, **121**(1-2):84-87.

3. Rochel N, Wurtz JM, Mitschler A, Klaholz B, Moras D: **The crystal structure of the nuclear receptor for vitamin D bound to its natural ligand**. *Mol Cell* 2000, **5**(1):173-179.

4. Yamada S, Yamamoto K, Masuno H: **Structure-function analysis of vitamin D and VDR model**. *Curr Pharm Design* 2000, **6**(7):733-748.

5. Yamamoto K, Masuno H, Choi M, Nakashima K, Taga T, Ooizumi H, Umesono K, Sicinska W, VanHooke J, DeLuca HF *et al*: **Three-dimensional modeling of and ligand docking to vitamin D receptor ligand binding domain**. *Proc Natl Acad Sci U S A* 2000, **97**(4):1467-1472.

6. Tocchini-Valentini G, Rochel N, Wurtz JM, Mitschler A, Moras D: **Crystal structures of the vitamin D receptor complexed to superagonist 20-epi ligands**. *Proc Natl Acad Sci U S A* 2001, **98**(10):5491-5496.

7. Tocchini-Valentini G, Rochel N, Wurtz JM, Moras D: **Crystal structures of the vitamin D nuclear receptor liganded with the vitamin D side chain analogues calcipotriol and seocalcitol, receptor agonists of clinical importance. Insights into a structural basis for the switching of calcipotriol to a receptor antagonist by further side chain modification**. *J Med Chem* 2004, **47**(8):1956-1961.

8. Hourai S, Fujishima T, Kittaka A, Suhara Y, Takayama H, Rochel N, Moras D: **Probing a water channel near the A-ring of receptor-bound 1 alpha,25-dihydroxyvitamin D3 with selected 2 alpha-substituted analogues**. *J Med Chem* 2006, **49**(17):5199-5205.

9. Yamada S, Yamamoto K: **Ligand recognition by vitamin D receptor: Total alanine scanning mutational analysis of the residues lining the ligand binding pocket of vitamin D receptor**. *Curr Top Med Chem* 2006, **6**(12):1255-1265.

10. Nakajima S, Hsieh JC, Jurutka PW, Galligan MA, Haussler CA, Whitfield GK, Haussler MR: **Examination of the potential functional role of conserved cysteine residues in the hormone binding domain of the human 1,25-dihydroxyvitamin D-3 receptor**. *J Biol Chem* 1996, **271**(9):5143-5149.

11. Solomon C, Macoritto M, Gao XL, White JH, Kremer R: **The unique tryptophan residue of the vitamin D receptor is critical for ligand binding and transcriptional activation**. *J Bone Miner Res* 2001, **16**(1):39-45.

12. Reddy MD, Stoynova L, Acevedo A, Collins ED: **Residues of the human nuclear vitamin D receptor that form hydrogen bonding interactions with the three hydroxyl groups of 1alpha,25-dihydroxyvitamin D3**. *J Steroid Biochem Mol Biol* 2007, **103**(3-5):347-351.

13. Swamy N, Xu WR, Paz N, Hsieh JC, Haussler MR, Maalouf GJ, Mohr SC, Ray R: **Molecular modeling, affinity labeling, and site-directed mutagenesis define the key points of interaction between the ligand-binding domain of the vitamin D nuclear receptor and 1 alpha,25-dihydroxyvitamin D_3**. *Biochemistry* 2000, **39**(40):12162-12171.

14. Azizi B, Chang EI, Doyle DF: **Chemical complementation: small-molecule-based genetic selection in yeast**. *Biochem Biophys Res Commun* 2003, **306**(3):774-780.

15. Schwimmer LJ, Rohatgi P, Azizi B, Seley KL, Doyle DF: **Creation and discovery of ligand-receptor pairs for transcriptional control with small molecules**. *Proc Natl Acad Sci U S A* 2004, **101**(41):14707-14712.

16. Taylor JL, Rohatgi P, Spencer HT, Doyle DF, Azizi B: **Characterization of a molecular switch system that regulates gene expression in mammalian cells through a small molecule**. *BMC Biotechnol*, **10**:12.

17. Baker K, Bleczinski C, Lin H, Salazar-Jimenez G, Sengupta D, Krane S, Cornish VW: **Chemical complementation: a reaction-independent genetic assay for enzyme catalysis**. *Proc Natl Acad Sci U S A* 2002, **99**(26):16537-16542.

18. James P, Halladay J, Craig EA: **Genomic libraries and a host strain designed for highly efficient two-hybrid selection in yeast**. *Genetics* 1996, **144**(4):1425-1436.

19. Jones G, Strugnell SA, DeLuca HF: **Current understanding of the molecular actions of vitamin D**. *Physiol Rev* 1998, **78**(4):1193-1231.

20. Neubig RR, Spedding M, Kenakin T, Christopoulos A: **International union of pharmacology committee on receptor nomenclature and drug classification. XXXVIII. Update on terms and symbols in quantitative pharmacology**. *Pharmacol Rev* 2003, **55**(4):597-606.

21. Mizwicki MT, Bula CM, Bishop JE, Norman AW: **New insights into Vitamin D sterol-VDR proteolysis, allostery, structure-function from the perspective of**

a conformational ensemble model. *J Steroid Biochem Mol Biol* 2007, **103**(3-5):243-262.

22. Struhl K, Davis RW: **Production of a functional eukaryotic enzyme in *Escherichia-coli* - cloning and expression of yeast structural gene for imidazoleglycerolphosphate dehydratase (HIS3)**. *Proc Natl Acad Sci U S A* 1977, **74**(12):5255-5259.

23. Castillo HS: **Mutational analysis and engineering of the human vitamin D receptor to bind a novel small molecule**. Atlanta: Georgia Institute of Technology; 2010.

24. Jurutka PW, Thompson PD, Whitfield GK, Eichhorst KR, Hall N, Dominguez CE, Hsieh JC, Haussler CA, Haussler MR: **Molecular and functional comparison of 1,25-dihydroxyvitamin D_3 and the novel vitamin D receptor ligand, lithocholic acid, in activating transcription of cytochrome P450 3A4**. *J Cell Biochem* 2005, **94**(5):917-943.

25. Mizwicki MT, Bula CM, Mahinthichaichan P, Henry HL, Ishizuka S, Norman AW: **On the mechanism underlying (23*S*)-25-dehydro-1alpha(OH)-vitamin D3-26,23-lactone antagonism of hVDRwt gene activation and its switch to a superagonist**. *J Biol Chem* 2009, **284**(52):36292-36301.

26. Ruau D, Duarte J, Ourjdal T, Perriere G, Laudet V, Robinson-Rechavi M: **Update of NUREBASE: nuclear hormone receptor functional genomics**. *Nucleic Acids Res* 2004, **32**:D165-D167.

27. Higuchi R, Krummel B, Saiki RK: **A general-method of invitro preparation and specific mutagenesis of DNA fragments - study of protein and DNA interactions**. *Nucleic Acids Res* 1988, **16**(15):7351-7367.

28. Trott O, Olson AJ: **AutoDock Vina: Improving the speed and accuracy of docking with a new scoring function, efficient optimization, and multithreading**. *J Comput Chem* 2009:456-461.

29. Sali A, Blundell TL: **Comparative protein modeling by satisfaction of spatial restraints**. *J Mol Biol* 1993, **234**(3):779-815.

30. Damborsky J, Prokop M, Koca J: **TRITON: graphic software for rational engineering of enzymes**. *Trends BiochemSci* 2001, **26**(1):71-73.

31. Pettersen EF, Goddard TD, Huang CC, Couch GS, Greenblatt DM, Meng EC, Ferrin TE: **UCSF chimera - A visualization system for exploratory research and analysis**. *J Comput Chem* 2004, **25**(13):1605-1612.

32. Sanner MF: **Python: A programming language for software integration and development**. *J Mol Graph* 1999, **17**(1):57-61.
33. Reichert J, Suhnel J: **The IMB Jena image library of biological macromolecules: 2002 update**. *Nucleic Acids Res* 2002, **30**(1):253-254.

34. Gietz RD, Woods RA: **Transformation of yeast by lithium acetate/single-stranded carrier DNA/polyethylene glycol method**. In: *Guide to Yeast Genetics and Molecular and Cell Biology, Pt B.* vol. 350. San Diego: Academic Press Inc; 2002: 87-96.

CHAPTER 3
ENGINEERING HUMAN VITAMIN D RECEPTOR TO BIND AND ACTIVATE IN RESPONSE TO CHOLECALCIFEROL

3.1 Random Mutagenesis

Cholecalciferol, an intermediate in the $1,25(OH)_2D_3$ biosynthetic pathway, was chosen as the target novel small molecule for protein engineering due to the similarities with $1,25(OH)_2D_3$, as well as the inactivity with wild-type hVDR. Cholecalciferol lacks the 1α and the 25-hydroxyl groups that are present in $1,25(OH)_2D_3$, which reduces hydrogen bonding potential between cholecalciferol and the surrounding residues. Due to cholecalciferol's similar structure to $1,25(OH)_2D_3$ and this ligands inability to activate wild-type hVDR, cholecalciferol was chosen as the target small molecule for hVDR engineering in an attempt to better understand the structure/function relationship between hVDR and the role of various residues in the ligand binding pocket. A determination of which residues are necessary to compensate for the lack of hydrogen bonding potential present in cholecalciferol could lead to a better understanding of the interactions that are necessary between hVDR and various ligands to obtain activation. In an attempt to obtain a hVDR variant that could bind cholecalciferol, random mutagenesis was performed using wild-type hVDR, H305F, and H305Y as starting templates based on their unique activation profiles, including the increased sensitivity of H305F and H305Y towards lithocholic acid.

Directed evolution is used in protein engineering to create mutants with desired functions that are nonexistent in nature. Directed evolution has emerged as a widely used method for protein engineering [1-4]. As a result, libraries of proteins can be created by rational or random mutagenesis [5, 6]. In rational design, structural analysis, mechanistic

and functional analyses of the targeted protein are required to make an informed decision regarding which residues to mutate in an attempt to engineer the protein for a certain function [7, 8]. However, random mutagenesis allows for a predominately unbiased library as well as a larger mutational spectrum (i.e. mutations not limited to the active site). Although rational design is a widely used method for protein engineering, a large amount of information is needed to accurately design a protein library, while random mutagenesis does not require the information that rational design needs. However, rational design does allow for more control over the variants that are produced than random mutagenesis

One method for random mutagenesis is DNA shuffling. In DNA shuffling, related sequences are digested by DNaseI to form a pool of DNA fragments [9, 10]. Mutations occur when a fragment from one gene anneals with a fragment from another gene. These recombinations create a library of chimeric genes, some of which may give the desired protein function [9, 10]. The staggered extension process (StEP), which increases the gene length over repeated cycles of PCR by annealing templates with complimentary ends is a method of random mutagenesis [11]. The extension process allows the primers to anneal to different templates and creating a library of chimeras [11]. Random drift mutagenesis involves the incorporation of neutral mutations into the proteins of interest [12].

Random mutagenesis, specifically error-prone PCR, was chosen over rational design to determine whether mutations beyond the four angstrom bias employed in the rational library design could lead to an ideal conformation, inducing activation of hVDR with cholecalciferol [13]. In this method, the gene of interest is subjected to a PCR reaction,

with a polymerase possessing a low fidelity (typically *Taq* polymerase) and lacking proofreading capability [14]. The addition of $MnCl_2$ further diminishes the polymerase's fidelity, as the manganese displaces the *Taq* polymerase's coordinating magnesium metal involved in stabilizing the enzyme [14-17]. Previously, error-prone PCR has been shown to be successful in the protein engineering of nuclear receptors for new receptor/ligand pairs, as was observed with the estrogen receptor and the synthetic ligand, 4,4'-dihydroxybenzil (DHB) [18-20].

3.2 Error-Prone Library

3.2.1 Experimental Techniques in Creating Library of Variants

A plasmid containing a Gal4DBD (GBD) fused to the hVDR variant, H305F (pGBDH305F), was used as the template DNA for error-prone PCR random mutagenesis. H305F was used as the template due to H305F's increased sensitivity towards LCA. Primers starting at the beginning of the gene and extending to the 3'UTR of the H305F gene were used to amplify the insert cassette. $MnCl_2$ was used to create a library of variants and the insert cassette and background vector were then co-transformed into the yeast strain PJ69-4A [21]. Via homologous recombination, the insert cassette and linearized background vector should yield a library of variants in the yeast. The transformants were plated onto adenine selective plates (SC-ALW) with LCA and cholecalciferol respectively, as well as non-selective plates (SC-LW) to determine transformation efficiency.

3.2.2 Results of Error-Prone Library

The variants were selected using chemical complementation, a system in which cholecalciferol binding a hVDR variant induces the recruitment of the coactivator/Gal4

activation domain fusion (GAD). The coactivator/GAD fusion will recruit the transcriptional machinery allowing the yeast to survive on media lacking adenine (SC-ALW).

This library displayed a transformation efficiency of 4.0×10^2 colony forming units/µg of vector cassette and a library size of 4.2×10^3 variants. Only one variant showed growth on adenine selective plates containing 10 µM cholecalciferol. This variant was tested for constitutive activity. As shown in Figure 3.1, growth was observed on the non-selective plate, as expected, since the non-selective plate selects for the two fusion plasmids. However, no growth was observed on the adenine selective plate without ligand, indicating that the variant is not constitutively active (Figure 3.1). The variant was rescued and sequenced, and sequencing showed that an additional mutation, H397Y, was added to the H305F template, creating the double variant H305F/H397Y.

To further characterize the variant, a liquid quantitation assay was performed. Yeast cells were added to adenine selective media with varying concentrations of LCA, cholecalciferol, and $1,25(OH)_2D_3$. H305F/H397Y showed activation at an EC_{50} value of 0.5 µM and 10-fold activation with LCA, as well as an EC_{50} value of >5 µM and 7-fold activation with cholecalciferol in adenine selective media (Figure 3.2). As shown in Figure 3.2A, Gal4, which is a ligand-independent transcription factor, grows as expected at all concentrations. H305F, H305Y, and H305F/H397Y display ligand-activated growth with LCA at an EC_{50} value of 0.5 µM and with an about 11-fold activation, whereas wild-type hVDR does not display growth with LCA in adenine selective media.

-LW **-ALW**

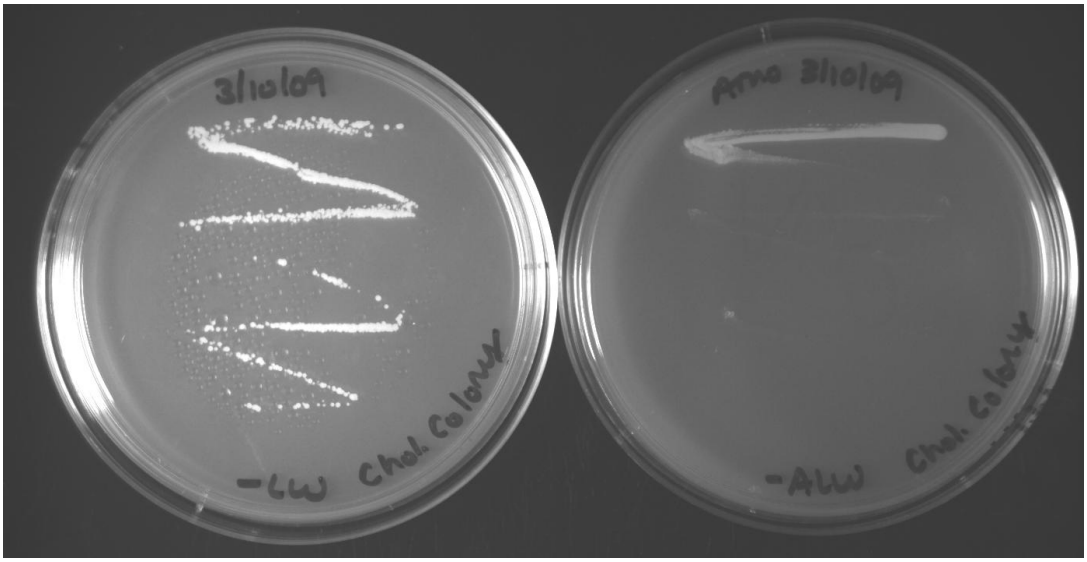

Figure 3.1: Streaking of Error Prone Selective colony. One selective colony grew on the selective plates. This colony was then streaked to test for constitutive activity. (A) Yeast growth should be observed on the non-selective plate (-LW). (B) Yeast growth is not observed on the adenine selective plate in the absence of ligand, implying that the variant is not constitutively active.

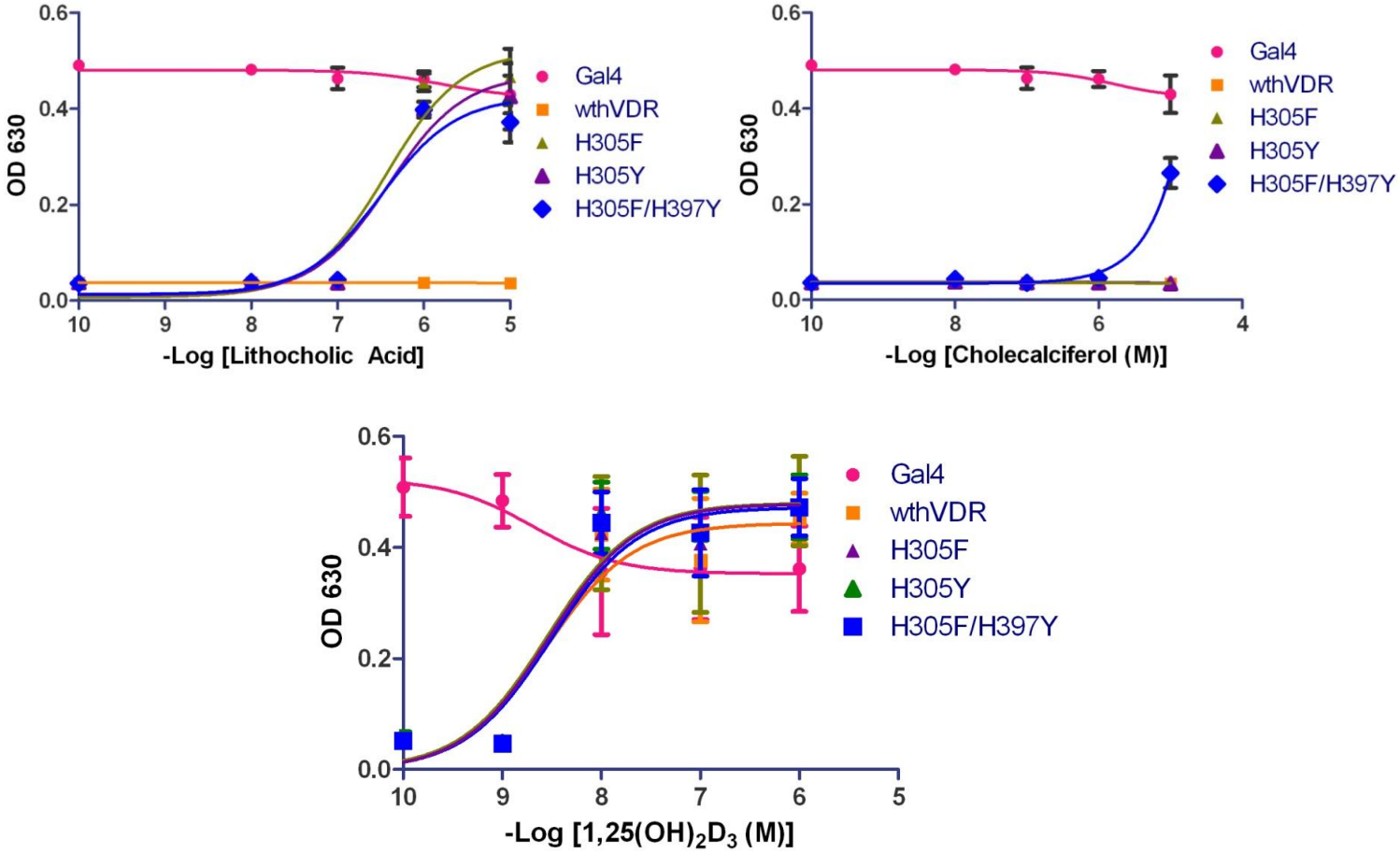

Figure 3.2: Testing variants with various ligands. (A) Ligand-activated growth is observed at 1 μM LCA for all variants. (B) Only H305F/H397Y display growth with cholecalciferol. (C) Wild-type hVDR and variants disply ligand-activated growth at 10 nM 1α,25-dihydroxyvitamin D_3.

H305F/H397Y is the only variant that displays ligand-activated growth with cholecalciferol with an EC_{50} value of 5 µM and a 7 fold-activation in adenine selective media (Figure 3.2B)

These results were surprising as previous mutational analysis had shown that H397 forms an essential hydrogen bond with the 25-hydroxyl group of $1,25(OH)_2D_3$, a bond not present in cholecalciferol. However, the combination of both mutations in the double variant results in an increase in bulk, since phenylalanine (189.9 Å) and tyrosine (196.3 Å) are bulkier than histidine (153.2 Å), hence decreasing the overall volume of the ligand binding pocket [22]. This suggests that a combination of volume changes as well as the presence of hydrogen bonding residues in the pocket at residues 305 and 397 leads to binding and activation with cholecalciferol, as seen with the H305F/H397Y mutant. With $1,25(OH)_2D_3$, H305F/H397Y displayed and EC_{50} value of 5 nM in adenine selective media with a 9-fold activation, which is the same as wild-type hVDR, H305F, and H305Y (Figure 3.2C).

3.2.3 Testing Variants in Mammalian Cell Culture

To determine whether the variants discovered with chemical complementation display the same or similar activity in mammalian cell assays, the variants were analyzed in human embryonic kidney 293T cells (HEK293T). HEK293T cells (ATCC, USA) were transfected with the following plasmids: pCMXwthVDR, pCMXH305F, pCMXH305Y, and pCMXH305F/H397Y. These plasmids contain the Gal4DBD (GBD) fused to the corresponding VDR ligand binding domain (GBD:LBD fusion under the control of a cytomegalovirus (CMV) promoter). The reporter plasmids used were p17*4TATAluc, containing the *Renilla* luciferase gene, and pCMXβgal, containing the β-

galactosidase gene, which allows for the determination of efficiency of the transfection. Lipofectamine 2000 served as the cationic lipid, which aids in the transmembrane transport of the plasmids. Ligands are added at various concentrations (0.01 μM- 100 μM LCA and 0.01 μM- 32 μM cholecalciferol), and cells are harvested and analyzed for luciferase and β-galactosidase activity, which allows for the calculation of transfection efficiency.

As shown in Figure 3.3, hVDR variants H305F and H305Y display activation at 3.3 μM LCA and 1 μM, respectively, in comparison to 100 μM LCA for wild-type hVDR as seen in yeast. H305F and H305Y show 40- and 74- fold activation, respectively, with LCA, whereas wild-type hVDR shows a 13-fold activation with LCA only at the highest concentration (100 μM) of ligand (Figure 3.3, Table 3.1). The H305F/H397Y variant has a 100-fold increase in sensitivity with LCA showing activation at 1 μM in comparison to 100 μM with wild-type hVDR and LCA. Additionally, H305F and H305Y display activation at 1 μM cholecalciferol resulting in a 54- and 65- fold activation, whereas no activation is observed with cholecalciferol and wild-type hVDR.

As can be seen in Table 3.1, EC_{50} values of 10 μM for LCA and 3 μM for cholecalciferol were observed with both H305F and H305Y, whereas, EC_{50} values for wild-type hVDR with both LCA and cholecalciferol were 100 μM or greater. The second generation double variant, H305F/H397Y, shows an EC_{50} value similar to that of H305F with LCA (10 μM), but an EC_{50} value of 300 nM is observed with cholecalciferol, in comparison to the 3 μM EC_{50} value for H305F. Thus, we have engineered a variant of hVDR that is now activated by submicromolar concentrations of cholecalciferol, along with an increased sensitivity to LCA.

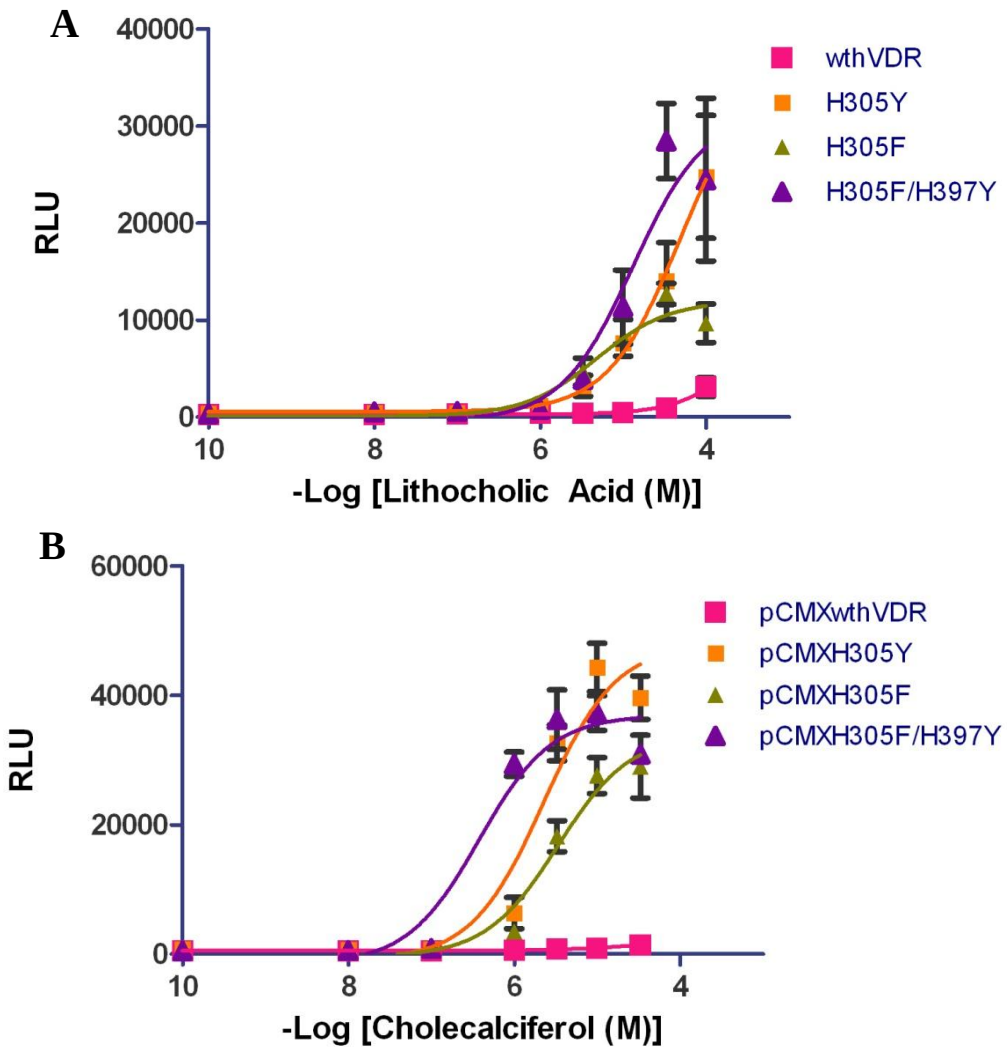

Figure 3.3: Testing variants in mammalian cell culture. Wild-type hVDR, H305F, H305Y, and H305F/H397Y were tested in cell culture. (A) Variants display activation with LCA beginning at 10 µM in comparison to 100 µM for wild-type. (B) H305F/H397Y displays a 10-fold increase in sensitivity towards cholecalciferol in comparison to H305F and H397Y.

Table 3.1: EC_{50} values and fold-activations of wthVDR and variants in mammalian cells (HEK293T).

	Lithocholic Acid		Cholecalciferol	
	HEK293T		HEK293T	
Variant	EC_{50}	FA	EC_{50}	FA
Wild-Type	100 μM	13 ± 7	>100 μM	3 ± 1
H305F	10 μM	40 ± 14	3 μM	54 ± 11
H305Y	10 μM	74 ± 25	3 μM	65 ± 6
H305F/H397Y	10 μM	84 ± 14	0.3 μM	70 ± 11

3.2.4 Modeling of Variants with Cholecalciferol

To visualize the impact of these mutations on ligand binding through *in silico* methods, cholecalciferol was docked into wild-type hVDR, H305F, and H305F/H397Y using AutoDock Vina, as described in the materials and methods (Figure 3.4) [23]. In the superimposition of wild-type hVDR and the variants with cholecalciferol, a drastic change is observed among these structures. Cholecalciferol, in the docked structures, shows a 180° rotation in the binding pocket when compared to the crystal structure of wild-type hVDR and $1,25(OH)_2D_3$. The 3-hydroxyl group of $1,25(OH)_2D_3$ has been shown to interact with S237 and R274, the 1α-hydroxyl group has been shown to interact with Y143 and S278, and the 25-hydroxyl group has been shown to interact with H305 and H397. Y143 has been implicated in ligand stabilization in the pocket, and the absence of the 1α-hydroxyl group allows the ligand to rotate. Previously Yamada *et al.* observed a 180° shift with 3-ketolithocholic acid in the hVDR pocket, similar to the observations with cholecalciferol (Figure 3.4A) [24].

In silico structures with cholecalciferol bound to wild-type hVDR (Figure 3.4A) also display a bent conformation of the ligand in comparison to the crystal structure of hVDR with $1,25(OH)_2D_3$, suggesting that a bowl-shaped conformation is preferred in order for the ligand to create the necessary contacts for binding and activation [25-27]. Therefore, the bent conformation with cholecalciferol and wild-type hVDR may be leading to an inferior conformation resulting from the absence of crucial contacts with R274, Y143, and S278 (Figure 3.4A). The absence of crucial contacts is most likely the cause for the lack of activation with wild-type hVDR and cholecalciferol.

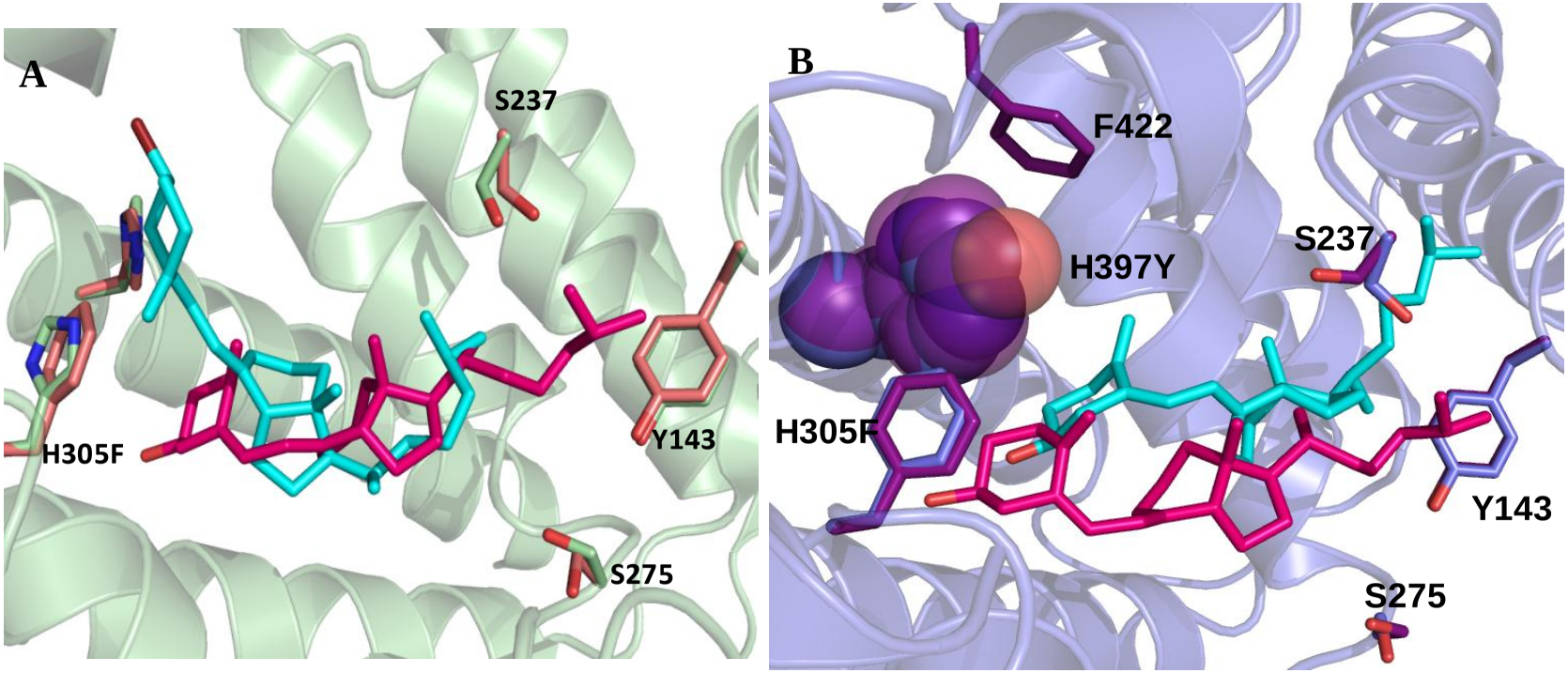

Figure 3.4: Docking of wthVDR, H305F, and H305F/H397Y with cholecalciferol. (A) wild-type = green (residues) and cyan (ligand), H305F= salmon (residues) and pink (ligand). (B) H305F= blue (residues) and pink (ligand), H305F/H397Y= purple (residues) and cyan (ligand).

In the case of the double variant H305F/H397Y, the combination of both mutations causes cholecalciferol to shift away from the tyrosine at residue 397. The presence of a tyrosine at position 397 instead of a histidine leads to a decreased binding pocket volume (Figure 3.4B). In the double variant, when comparing the spacefill model of histidine and tyrosine at 397, the bulkier tyrosine results in an upward shift of the cholecalciferol in the pocket, perhaps allowing for more favorable van der Waals interactions between cholecalciferol and the residues in the pocket. In the docked structures, the hydroxyl group of cholecalciferol is unlikely to be within hydrogen bonding distance of the histidine (6.09 Å) or tyrosine (6.59 Å) of the variants. However, this could be an artifact of the docking program as the protein remains static during docking; therefore, the residues may be in a different conformation and may be able to form a hydrogen bond with cholecalciferol. In addition, water molecules are known to form hydrogen bonds with ligands, such as in the case of rat VDR [24, 28]. Thus, a water molecule may be present in the ligand binding pocket capable, of forming a hydrogen bond with H397Y and cholecalciferol [24]. Thus, the increase in bulk at 397 also increases the favorable hydrophobic van der Waals interactions between residue 397 and F422, an important residue in the AF-2 domain of hVDR. This may be leading to a stabilizing effect, such as tighter co-activator binding, as shown previously [29, 30]. The H397-F422 bridge creates a 19 Å distance between E420 on the AF-2 and K246 of helix 3. This charge clamp is essential in creating the CoAc binding site on wild-type hVDR [29].

3.3 Summary

When H305F was subjected to random mutagenesis using error-prone PCR in order to discover a variant capable of binding cholecalciferol, a double variant, H305F/H397Y, displayed LCA activation profiles similar to those of H305F but a ~10-fold increased sensitivity towards cholecalciferol, displaying an EC_{50} of 300 nM in mammalian cells (Figure 3.2, 3.3). *In silico* docking of cholecalciferol to the H305F variant using AutoDock Vina displays a less linear conformation for cholecalciferol in comparison to the same ligand bound to wild-type hVDR (Figure 3.4A) [23].

In conclusion, more insight into the relationship and crucial interactions between hVDR and various ligands has been acquired. These results have shown that hydrogen bonding at H397 is necessary for binding and activation with cholecalciferol. Also, the 180° flip of cholecalciferol in the ligand binding pocket suggests that the hydrogen bonding potential at the H305/H397 end of the pocket is also important for stabilizing cholecalciferol and allowing for ligand activation. Single point mutations at position 305, H305F and H305Y, result in an increased sensitivity with LCA and cholecalciferol. A hVDR variant, H305F/H397Y, is able to bind cholecalciferol at submicromolar concentrations in mammalian cells, whereas wild-type hVDR does not display activation with this ligand. Additionally, the H397Y mutation in combination with H305F, introduces a modification in the volume of the pocket, leading to activation with cholecalciferol. Future work, including *in vitro* binding studies, as well as additional mutational analysis, to determine if the tyrosine is necessary for maximal activation with cholecalciferol will be performed. This will allow a more comprehensive understanding of the dynamic interactions between wild-type hVDR and these variants with their potential ligands.

3.4 Materials and Methods

3.4.1 Ligands

Lithocholic acid, cholecalciferol, and 1α,25-dihydroxyvitamin D_3 were ordered from MP Biomedicals, LLC (Solon, OH), Sigma (St. Louis, MO), and BIOMOL (Plymouth Meeting, PA), respectively. 10 mM stocks of lithocholic acid and cholecalciferol and a 13.3 μM stock of 1α,25-dihydroxyvitamin D_3 were made with 80% ethanol: 20% DMSO and stored at 4 °C.

3.4.2 Error-Prone PCR

pGBDhVDRH305F, the plasmid containing the hVDR mutant, was used as the template DNA, for error-prone PCR random mutagenesis with primers:
5'-gatcctgaagcggaaggagg-3' and
5'-gtccaggcagggtggccagaacgggtgggcacaaaggatggactagttcaggagatctcattgccaaacacttcgag-3'. 0.5 μM of each primer, 500 μM dNTPs, 7 mM $MgCl_2$, 20 μM $MnCl_2$, 250 ng template DNA (pGBDhVDRH305F), 1x *Taq* buffer, and 5 units of *Taq* polymerase (Fermentas, USA) were used for each reaction. The PCR program used was: 95 ºC for 5 min, 95 ºC for 30 sec, 55 ºC for 30 sec, 72 ºC for 60 sec, 20 cycles yielding ~ 1 μg of variant VDR insert cassette. The insert cassettes were then transformed into yeast using the 1x TRAFCO protocol and tested in liquid quantitation assays using chemical complementation as mentioned in Schwimmer *et al.* [31].

3.4.3 Yeast Transformation

Using the 1x TRAFCO yeast transformation protocol [32], 100 ng of the digested background plasmid (pGBDhVDRbackground) and 900 ng of insert cassette were transformed into the yeast strain PJ69-4A. A plasmid containing a fusion of the human

co-activator ACTR and the Gal4 activation domain with a leucine marker was also co-transformed [21]. Transformants were plated onto synthetic complete (SC) agar plates lacking adenine, leucine, and tryptophan (SC-ALW) with various ligands. Transformants were also plated onto synthetic complete agar plates lacking leucine and tryptophan (SC-LW) containing no ligand, to determine transformation efficiency. Transformation efficiency was calculated based on the number of colonies on the plate divided by the µg of vector cassette that was plated onto each plate.

3.4.4 Liquid Quantitation Assay in Yeast

Variants were tested in liquid quantitation assays in 96-well plates with media lacking adenine, leucine, and tryptophan (SC-ALW), with or without LCA, cholecalciferol, or $1,25(OH)_2D_3$ at varying concentrations (ranging from 10 nM - 10 µM for LCA and cholecalciferol, 1 nM - 1 µM for $1,25(OH)_2D_3$). A 4:1 ratio of media (SC-ALW): cells (yeast resuspended in water) were aliquoted into 96-well plates. Plates were incubated at 30°C, with shaking at 170 rpm. Optical density (OD) readings at 630 nm were recorded at 0, 24, and 48 hours as a measure of growth density.

3.4.5 Mammalian Cell Culture

HEK293T cells (ATCC, USA) were transfected with the following plasmids: pCMXwild-type hVDR, pCMXH305F, pCMXH305Y, and pCMXH305F/H397Y. These plasmids contain the Gal4DBD (GBD) fused to the corresponding VDR ligand binding domain (GBD:LBD fusion under the control of a cytomegalovirus (CMV) promoter). The reporter plasmids were p17*4TATAluc, containing the *Renilla* luciferase gene under the control of four Gal4 response elements located upstream from a minimal thymidine kinase promoter, and pCMXβgal, a plasmid containing the β-galactosidase gene under

the control of the mammalian CMV promoter. Lipofectamine 2000 (Invitrogen, USA) served as the cationic lipid and transfection experimental details are described in Taylor *et al*. [33]. The ligands were added to the wells at various concentrations ((0.01 μM- 100 μM) LCA and (0.01 μM- 32 μM) cholecalciferol). Cells were harvested and analyzed for luciferase and β-galactosidase activity. All data points represent the average of triplicate experiments normalized against β-galactosidase activity. Error bars represent the standard deviation calculated using standard deviation: σ = Square root ($\Sigma[(X-\mu)^2]/N$). Fold activation was calculated by dividing the value at maximal activation by the value at the no ligand data point.

3.4.6 Docking hVDR Variants

The structures of the hVDR mutants were prepared *in silico* using the program TRITON 4.0.0 (National Centre for Biomolecular Research, Czech Republic) and its external program MODELLER (National Centre for Biomolecular Research, Czech Republic). These programs do not include any molecular dynamics. The computational site directed mutagenesis method, which is based on using the wild-type protein for homology modeling was employed [34, 35]. The wild-type hVDRLBD from the crystal structure 1DB1 was used as the template, which is missing residues 165-215 for crystallization purposes [36]. These residues were not added computationally in the modeling process.

Variants were prepared for docking using the UCSF CHIMERA- interactive molecular graphics program by: (1) removing the ligand and water molecules, (2) adding polar hydrogens, and (3) assigning Gasteiger charges [37]. Ligands were created using ChemBioDraw Ultra 11.0 and ChemBio3D Ultra 11.0 (Cambridge Soft, USA) and

modified with the AutoDockTools by adding Gasteiger charges, setting the partial charge property of each atom [38]. AutoDock Vina was used to perform docking simulations with default parameters [23]. In the Autodock simulations the hVDR ligand binding domain was held rigid while the ligand was allowed to rotate, based on quantum mechanical rotations that are incorporated into the algorithm. For the ligands the C3-OH bond, the carbon chain connecting the two ring systems, and the aliphatic chain extending off of C17 were allowed to rotate freely in the simulations. The solutions with the lowest free energy of binding were analyzed.

3.5 References

1. Tao H, Cornish VW: **Milestones in directed enzyme evolution**. *Curr Opin Chem Biol* 2002, **6**(6):858-864.

2. Roodveldt C, Aharoni A, Tawfik DS: **Directed evolution of proteins for heterologous expression and stability**. *Curr Opin Struct Biol* 2005, **15**(1):50-56.

3. Schmidt-Dannert C: **Directed evolution of single proteins, metabolic pathways, and viruses**. *Biochemistry* 2001, **40**(44):13125-13136.

4. Williams GJ, Nelson AS, Berry A: **Directed evolution of enzymes for biocatalysis and the life sciences**. *Cell Mol Life Sci* 2004, **61**(24):3034-3046.

5. Arnold FH: **When blind is better: protein design by evolution**. *Nat Biotechnol* 1998, **16**(7):617-618.

6. Bornscheuer UT, Pohl M: **Improved biocatalysts by directed evolution and rational protein design**. *Curr Opin Chem Biol* 2001, **5**(2):137-143.

7. Voigt CA, Mayo SL, Arnold FH, Wang ZG: **Computationally focusing the directed evolution of proteins**. *J Cell Biochem Suppl* 2001, **Suppl 37**:58-63.

8. Voigt CA, Mayo SL, Arnold FH, Wang ZG: **Computational method to reduce the search space for directed protein evolution**. *Proc Natl Acad Sci U S A* 2001, **98**(7):3778-3783.

9. Stemmer WP: **Rapid evolution of a protein in vitro by DNA shuffling**. *Nature* 1994, **370**(6488):389-391.

10. Stemmer WP: **DNA shuffling by random fragmentation and reassembly: in vitro recombination for molecular evolution**. *Proc Natl Acad Sci U S A* 1994, **91**(22):10747-10751.

11. Vanhercke T, Ampe C, Tirry L, Denolf P: **Reducing mutational bias in random protein libraries**. *Anal Biochem* 2005, **339**(1):9-14.

12. Bergquist PL, Reeves RA, Gibbs MD: **Degenerate oligonucleotide gene shuffling (DOGS) and random drift mutagenesis (RNDM): two complementary techniques for enzyme evolution**. *Biomol Eng* 2005, **22**(1-3):63-72.

13. Bloom JD, Arnold FH: **In the light of directed evolution: Pathways of adaptive protein evolution**. *Proc Natl Acad Sci U S A* 2009, **106**:9995-10000.

14. Pritchard L, Corne D, Kell D, Rowland J, Winson M: **A general model of error-prone PCR**. *J Theor Biol* 2005, **234**(4):497-509.

15. Moore GL, Maranas CD: **Modeling DNA mutation and recombination for directed evolution experiments**. *J Theor Biol* 2000, **205**(3):483-503.

16. Moore JC, Jin HM, Kuchner O, Arnold FH: **Strategies for the in vitro evolution of protein function: enzyme evolution by random recombination of improved sequences**. *J Mol Biol* 1997, **272**(3):336-347.

17. Zhao H, Arnold FH: **Combinatorial protein design: strategies for screening protein libraries**. *Curr Opin Struct Biol* 1997, **7**(4):480-485.

18. Islam KM, Dilcher M, Thurow C, Vock C, Krimmelbein IK, Tietze LF, Gonzalez V, Zhao H, Gatz C: **Directed evolution of estrogen receptor proteins with altered ligand-binding specificities**. *Protein Eng Des Sel* 2009, **22**(1):45-52.

19. McLachlan MJ, Chockalingam K, Lai KC, Zhao H: **Directed evolution of orthogonal ligand specificity in a single scaffold**. *Angew Chem Int Ed Engl* 2009, **48**(42):7783-7786.

20. Chockalingam K, Chen Z, Katzenellenbogen JA, Zhao H: **Directed evolution of specific receptor-ligand pairs for use in the creation of gene switches**. *Proc Natl Acad Sci U S A* 2005, **102**(16):5691-5696.

21. Azizi B, Chang EI, Doyle DF: **Chemical complementation: small-molecule-based genetic selection in yeast**. *Biochem Biophys Res Commun* 2003, **306**(3):774-780.

22. Reichert J, Suhnel J: **The IMB Jena image library of biological macromolecules: 2002 update**. *Nucleic Acids Res* 2002, **30**(1):253-254.

23. Trott O, Olson AJ: **AutoDock Vina: Improving the speed and accuracy of docking with a new scoring function, efficient optimization, and multithreading**. *J Comput Chem* 2009:456-461.

24. Yamada S, Yamamoto K: **Ligand recognition by vitamin D receptor: Total alanine scanning mutational analysis of the residues lining the ligand binding pocket of vitamin D receptor**. *Curr Top Med Chem* 2006, **6**(12):1255-1265.

25. Tocchini-Valentini G, Rochel N, Wurtz JM, Mitschler A, Moras D: **Crystal structures of the vitamin D receptor complexed to superagonist 20-epi ligands**. *Proc Natl Acad Sci U S A* 2001, **98**(10):5491-5496.

26. Tocchini-Valentini G, Rochel N, Wurtz JM, Moras D: **Crystal structures of the vitamin D nuclear receptor liganded with the vitamin D side chain analogues calcipotriol and seocalcitol, receptor agonists of clinical importance. Insights into a structural basis for the switching of calcipotriol to a receptor antagonist by further side chain modification**. *J Med Chem* 2004, **47**(8):1956-1961.

27. Hourai S, Fujishima T, Kittaka A, Suhara Y, Takayama H, Rochel N, Moras D: **Probing a water channel near the A-ring of receptor-bound 1 alpha,25-dihydroxyvitamin D3 with selected 2 alpha-substituted analogues**. *J Med Chem* 2006, **49**(17):5199-5205.

28. Vanhooke JL, Benning MM, Bauer CB, Pike JW, DeLuca HF: **Molecular structure of the rat vitamin D receptor ligand binding domain complexed with 2-carbon-substituted vitamin D3 hormone analogues and a LXXLL-containing coactivator peptide**. *Biochemistry* 2004, **43**(14):4101-4110.

29. Carlberg C: **Molecular basis of the selective activity of vitamin D analogues**. *J Cell Biochem* 2003, **88**(2):274-281.

30. Mizwicki MT, Bula CM, Bishop JE, Norman AW: **New insights into Vitamin D sterol-VDR proteolysis, allostery, structure-function from the perspective of a conformational ensemble model**. *J Steroid Biochem Mol Biol* 2007, **103**(3-5):243-262.

31. Schwimmer LJ, Rohatgi P, Azizi B, Seley KL, Doyle DF: **Creation and discovery of ligand-receptor pairs for transcriptional control with small molecules**. *Proc Natl Acad Sci U S A* 2004, **101**(41):14707-14712.

32. Gietz RD, Woods RA: **Transformation of yeast by lithium acetate/single-stranded carrier DNA/polyethylene glycol method**. In: *Guide to Yeast Genetics and Molecular and Cell Biology, Pt B.* vol. 350. San Diego: Academic Press Inc; 2002: 87-96.

33. Taylor JL, Rohatgi P, Spencer HT, Doyle DF, Azizi B: **Characterization of a molecular switch system that regulates gene expression in mammalian cells through a small molecule**. *BMC Biotechnol*, **10**:12.

34. Sali A, Blundell TL: **Comparative protein modeling by satisfaction of spatial restraints**. *J Mol Biol* 1993, **234**(3):779-815.

35. Damborsky J, Prokop M, Koca J: **TRITON: graphic software for rational engineering of enzymes**. *Trends BiochemSci* 2001, **26**(1):71-73.

36. Rochel N, Wurtz JM, Mitschler A, Klaholz B, Moras D: **The crystal structure of the nuclear receptor for vitamin D bound to its natural ligand**. *Mol Cell* 2000, **5**(1):173-179.

37. Pettersen EF, Goddard TD, Huang CC, Couch GS, Greenblatt DM, Meng EC, Ferrin TE: **UCSF chimera - A visualization system for exploratory research and analysis**. *J Comput Chem* 2004, **25**(13):1605-1612.

38. Sanner MF: **Python: A programming language for software integration and development**. *J Mol Graph* 1999, **17**(1):57-61.

CHAPTER 4
INVESTIGATING THE ROLE OF H305 AND H397 IN HVDR IN LIGAND ACTIVATION

4.1 Introduction

As shown in Chapters 2 and 3, three variants, H305F, H305Y, and H305F/H397Y were discovered to have a 100-fold increased sensitivity towards lithocholic acid in comparison to wild-type, where an EC_{50} of 10 μM and fold activations of 40 $\pm$ 14 and 74 $\pm$ 25 were observed for H305F and H305Y, respectively in mammalian cells. H305F/H397Y, displays a 10-fold increase in sensitivity with cholecalciferol, with an EC_{50} of 0.3 μM and a 70 $\pm$ 11 fold activation was observed, whereas wild-type hVDR shows no activation with cholecalciferol in mammalian cells. Due to these results, lithocholic acid and cholecalciferol were modeled using Auto Dock Vina into the ligand binding pocket of wild-type hVDR in an attempt to visualize the most obvious differences between these ligands in the pocket in comparison to $1,25(OH)_2D_3$ [1]. As can be seen in Figure 4.1A, lithocholic acid does not occupy as much space in the LBP when compared to $1,25(OH)_2D_3$. Therefore, decreasing the volume of the LBP should allow for increased sensitivity of the receptor towards LCA in comparison to wild-type hVDR by creating closer contacts with the residues in the pocket. The structure of cholecalciferol appears bent when compared to the natural bowl shape of $1,25(OH)_2D_3$ in the crystal structure [2].

Previously, Mizwicki *et al.* showed that H305F and H305F/H397F caused the wild-type hVDR antagonist, MK, to act as a superagonist. Their argument is that MK occupies less space in the hVDR ligand binding pocket and therefore does not make the necessary contacts with H305 and H397. This hypothesis is based upon the fact that

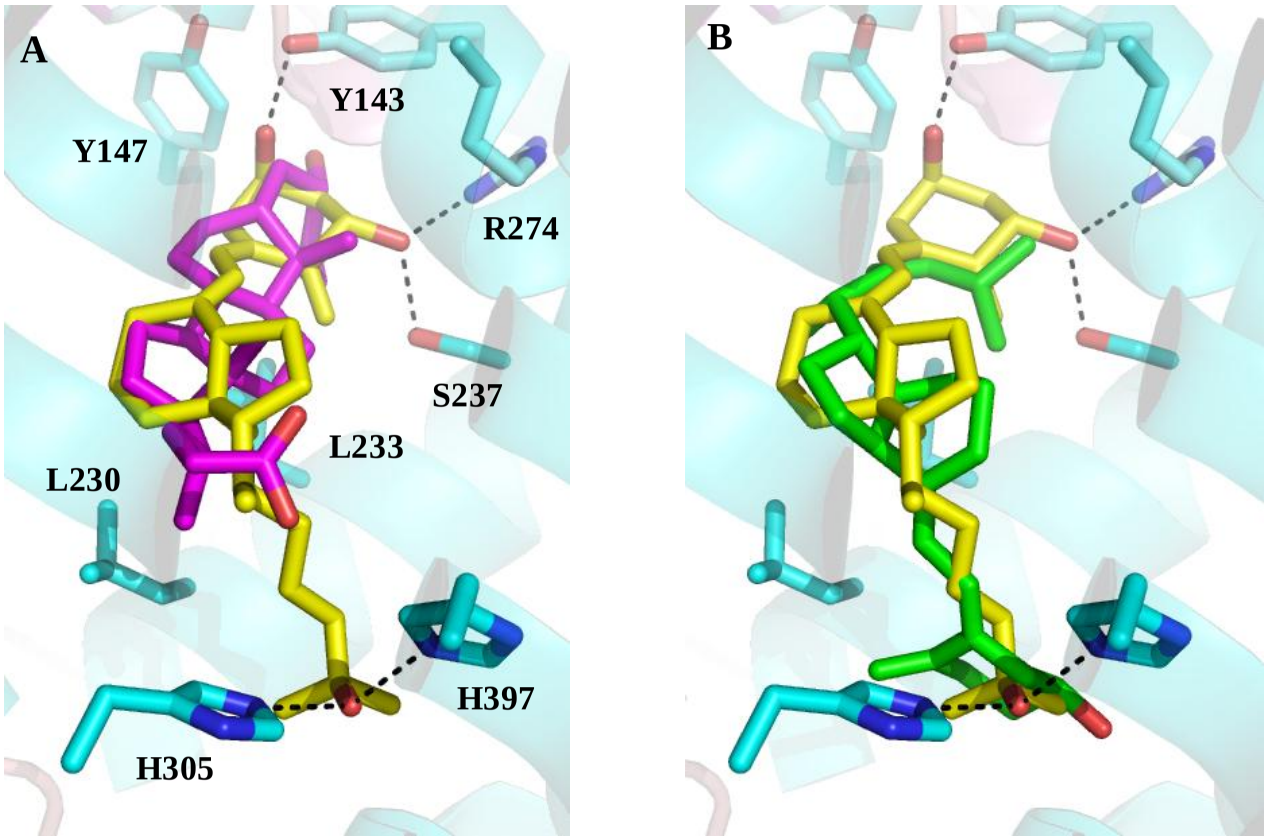

Figure 4.1: Lithocholic Acid and Cholecalciferol from docking overlayed with $1,25(OH)_2D_3$ from the crystal structure 1DB1. (A)Yellow is $1,25(OH)_2D_3$ and Lithocholic Acid is magenta. (B) Cholecalciferol is green. This structure represents the various placements of the ligands in the binding pocket in comparison to $1,25(OH)_2D_3$. As can be seen in A, LCA appears to take up less space in the LBP. Also, cholecalciferol is slightly bent in comparison to $1,25(OH)_2D_3$.

when H305 and H397 are mutated to residues bulkier than histidine, the volume of the LBP decreases, enhancing the interactions between MK and both H305 and H397, and increasing van der Waals contacts with surrounding residues, such as F401 and F422. This could then lead to tighter coactivator binding and increased activation [3, 4].

Based upon the results from Chapters 2 and 3, the idea was developed such that in order to obtain maximal activation with cholecalciferol, a combination of a bulky residue and a bulky residue with hydrogen bonding potential was needed. Based on the results with H305F/H397Y and the previous findings of Mizwicki *et al.*, a series of single and double mutants at the H305 and H397 positions with both phenylalanine and tyrosine were created in an effort to determine whether increased sensitivity would also be observed with H397F and H397Y, as was observed with H305F and H305Y. Due to H397's location on the same side of the pocket as H305, a variation in the activation profile should be noticed. The double variants were created in an attempt to determine the combination of residues required for maximal activation as well as increased sensitivity with cholecalciferol.

4.2 Phenylalanine and Tyrosine Variants

4.2.1 Testing Phenylalanine and Tyrosine Variants in Yeast

Wild-type hVDR was subjected to site-directed mutagenesis in order to create a combination of single and double variants with phenylalanine and tyrosine. The variants H397F, H397Y, H305F/H397F, H305Y/H397F, and H305Y/H397Y were created via site-directed mutagenesis. These variants in addition to H305F, H305Y, H305F/H397Y, wild-type hVDR, and Gal4 were tested with $1,25(OH)_2D_3$, LCA, and cholecalciferol.

Gal4 was used as the positive control, and wild-type hVDR was used as the negative control.

The variants were first tested in chemical complementation with $1,25(OH)_2D_3$ to determine whether the mutations would alter the activation profile observed with the wild-type ligand, $1,25(OH)_2D_3$. Chemical complementation links the activation of a nuclear receptor by a small molecule to the survival of yeast. Therefore, if the ligand activates the variant, then yeast growth should be observed. As can be seen in Figure 4.2A, all variants including wild-type showed ligand-activated growth with $1,25(OH)_2D_3$ at and EC_{50} of 5 nM with about 10-fold activation with the exception of H397F and H305F/H397F. Ligand-activated growth was not observed with H397F. The double variant, H305F/H397F, was observed to be constitutively active (growth is observed in the absence of ligand), confirming the hypothesis made by Mizwicki *et al.* [3]. Using molecular dynamics, a series of calculations showed that H305F/H397F should remain in the active conformation in the absence of a ligand, causing a constitutively active receptor. Although H305F/H397F was later shown to be ligand-activated by Mizwicki *et al.*, these results confirm their original hypothesis [3, 4].

The variants were tested for growth in yeast with LCA to determine whether a trend could be observed relating activation profiles and mutations. hVDR is hypothesized to require a bulky residue at H305 in order to obtain activation with LCA in yeast. The only variants that should display increased sensitivity are H305F, H305Y, H305F/H397Y, H305Y/H397F, and H305Y/H397Y due to the presence of a bulky residue at H305. As shown in Figure 4.2B the hypothesis was confirmed, such that only the variants that

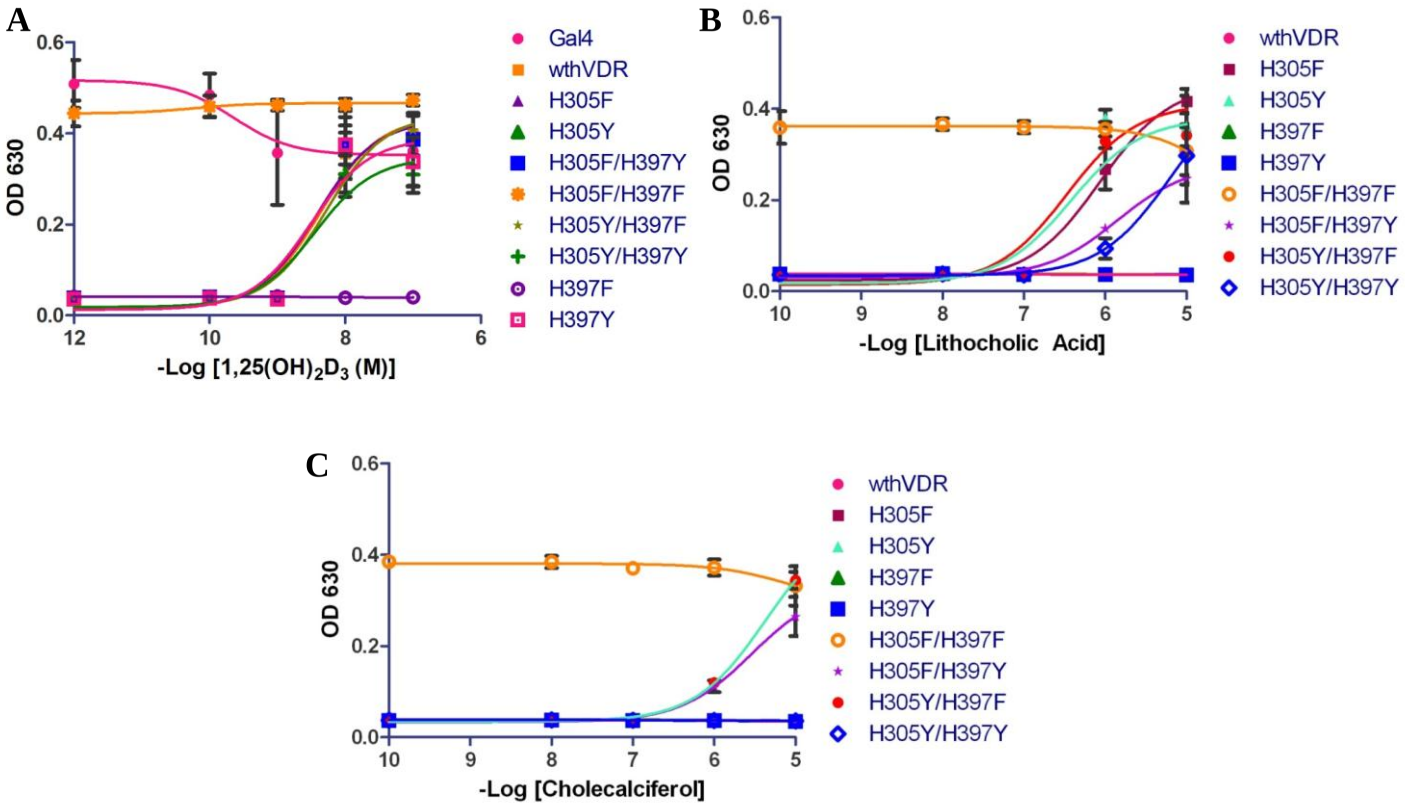

Figure 4.2: Testing hVDR Site-directed variants with various ligands in yeast. (A) All variants and wild-type hVDR display ligand-activated growth at 10 nM $1,25(OH)_2D_3$ in adenine selective media with the exception of H397F and H305F/H397F. (B) All variants display enhanced sensitivity towards LCA in adenine selective media, except H397F, H397Y, and wild-type hVDR. (C) Only H305F/H397Y and H305Y/H397F display ligand-activated growth with cholecalciferol in adenine selective media.

contained a mutation at H305 displayed ligand-activated growth in response to LCA with the exception of H305F/H397F. H305F, H305Y, H305F/H397Y, H305Y/H397F, and H305Y/H397Y display an EC_{50} of 0.5 μM with an 8-fold activation in adenine selective media. These results are in comparison to wild-type, H397F, and H397Y, which do not show ligand-activated growth with LCA in adenine selective media, as well as H305F/H397F, which is constitutively active.

As described in Chapter 3, the double variant, H305F/H397Y, shows ligand-activated growth with cholecalciferol at 10 μM in yeast. The hypothesis is that the combination of a bulky residue and a bulky residue with hydrogen bonding potential, as seen with phenylalanine and tyrosine, is needed for activation with cholecalciferol in yeast. When the variants were tested with cholecalciferol, only H305F/H397Y and H305Y/H397F show ligand-activated growth in adenine selective media displaying an EC_{50} of 1 μM cholecalciferol with a 8-fold activation, the other variants, including wild-type hVDR, show no ligand-activated growth. H305F/H397F was constitutively active (Figure 4.2C). These results confirm the hypothesis that a combination of a bulky residue and a bulky residue with hydrogen bonding potential is needed at H305 and H397, respectively, in order to obtain ligand activation with cholecalciferol in yeast.

All of the variants display wild-type activation levels with $1,25(OH)_2D_3$, except H397F and H397Y, which suggests that hydrogen bonding is not essential at H305. However, the results do suggest that H397 is less tolerant to mutations than H305. Only variants containing a mutation at H305 showed increased sensitivity with LCA. This result confirms the hypothesis that bulk is needed at H305 in order to obtain activation with LCA. As for cholecalciferol, only the two variants that contained a combination of

a bulky residue and bulky residue with hydrogen bonding potential at H305 and H397, H305F/H397Y and H305Y/H397F, displayed ligand-activated growth with cholecalciferol, as expected. This suggests that the presence of histidines in wild-type hVDR at 305 and 397 ensures VDR binding and activation in response to $1,25(OH)_2D_3$ and not cholecalciferol. This could potentially be due to the strictly regulated production of $1,25(OH)_2D_3$. If wild-type hVDR could slightly bind cholecalciferol, regulation would be lost, causing a drastic effect on the regulation of calcium homeostasis, as well as, other regulatory pathways [5].

4.2.2 Testing Phenylalanine and Tyrosine Variants in Mammalian Cells

To confirm the results that were observed in yeast, the variants were cloned into a mammalian expression vector, pCMXGBDhVDR, via site-directed mutagenesis in order to be tested in mammalian cell culture. pCMXGBDhVDR, a mammalian expression vector, contains the Gal4 DNA binding domain fused to the hVDR gene under the control of a cytomegalovirus promoter. The variants were tested with all three ligands ($1,25(OH)_2D_3$, LCA, and cholecalciferol). The results with $1,25(OH)_2D_3$ were confirmed, as all of the variants in addition to wild-type showed activation at an EC_{50} of 3 nM except for H305F/H397F, which was constitutively active displaying a fold activation of 3 $\pm$2 (Table 4.1, Figure 4.3A). The only observed difference between mammalian cells and yeast was regarding H397F, which did not display ligand-activated growth in yeast but displays an EC_{50} of 3 nM in mammalian cells. This is most likely due to the differences between the two systems. In chemical complementation a specific coactivator is introduced into the system, whereas mammalian cells express an array of coactivators that can be recruited for transcriptional activation. Based on the results of

Mizwicki *et al.*, hydrogen bonding residues should not be necessary at H305 and H397 to achieve activation with $1,25(OH)_2D_3$ [3, 4]. This hypothesis was confirmed by the fact that all of the variants, whether containing a phenylalanine or tyrosine, displayed wild-type activation levels with $1,25(OH)_2D_3$ in mammalian cells (Figure 4.3A).

The variants were then tested in mammalian cell culture with LCA to determine whether the same trends were observed as in yeast. Only variants with a mutation at H305 displayed an increased sensitivity similar to yeast towards LCA in comparison to wild-type hVDR (Figure 4.3B). All site-directed variants containing a mutation at position 305 showed activation with LCA, whereas the two single variants containing a mutation at position 397, H397F and H397Y, showed no activation with LCA. Wild-type hVDR only showed slight activation at the highest concentration. An EC_{50} of 10 μM was observed for H305F, H305Y, H305F/H397Y, and H305Y/H397F, with fold activations of 46 $\pm$ 22, 78 $\pm$ 15, 50 $\pm$ 3, and 47 $\pm$ 2, respectively. EC_{50}'s of 50 μM and greater than 100 μM with fold activations of 69 $\pm$ 9 and 10 $\pm$ 5 were observed for H305Y/H397Y and wild- type hVDR, respectively (Table 4.1). H305FH397F was also shown to be constitutively active in mammalian cells as well as yeast displaying a 3 $\pm$ 2 fold activation. Considering these results with LCA, the hypothesis was that a bulky residue was needed at H305 to increase sensitivity. This was confirmed by the fact that only variants containing a mutation at H305 (H305F, H305Y, H305F/H397Y, H305Y/H397F, and H305Y/H397Y) displayed increased activation with LCA in yeast and mammalian cells (Figure 4.2B, Figure 4.3B).

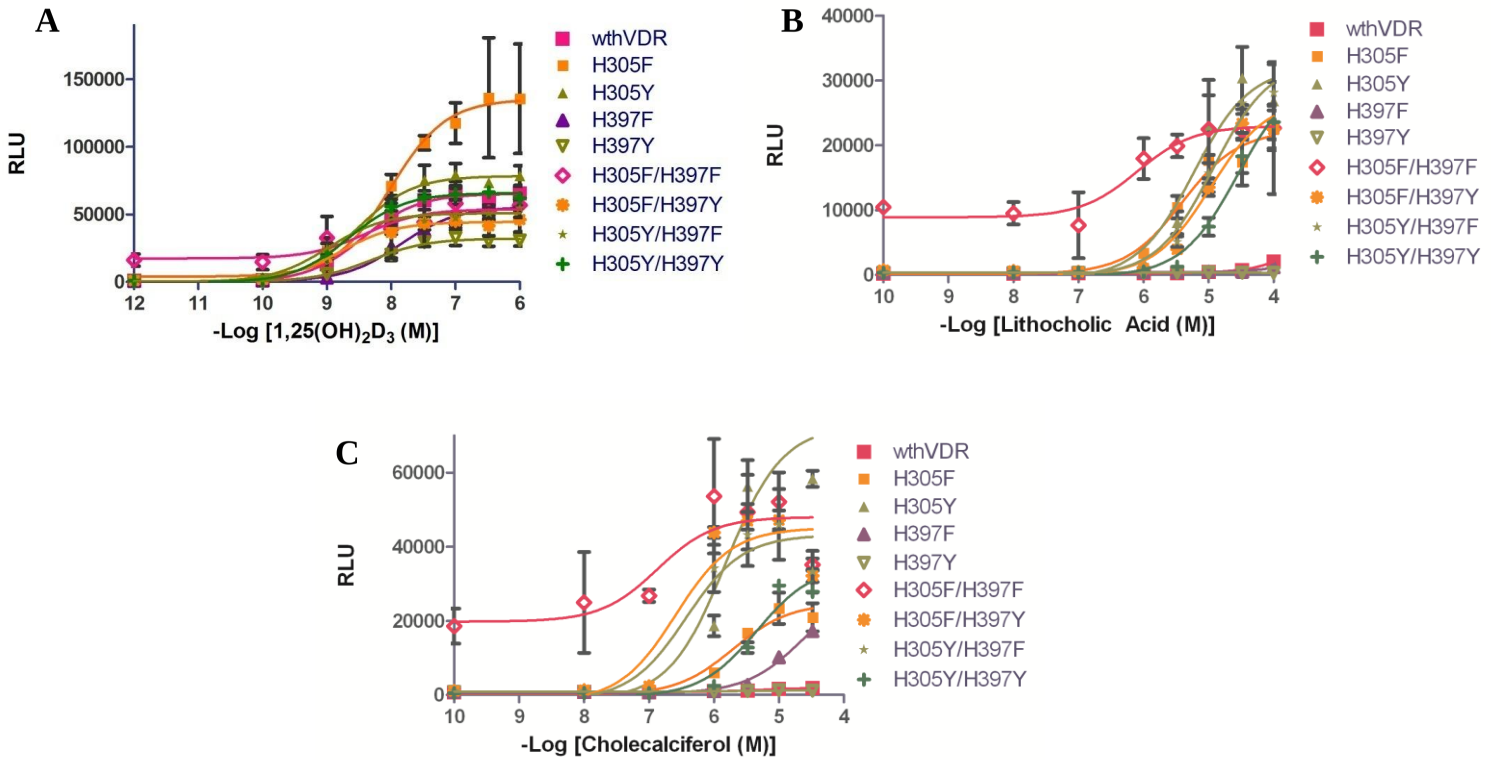

Figure 4.3: Testing hVDR Site-directed variants with various ligands in mammalian cells. A) All variants in addition to wild-type hVDR display activation with $1,25(OH)_2D_3$. B) All variants display increased sensitivity towards LCA in comparison to wild-type hVDR, except H397F and H397Y. C) All of the variants display activation with cholecalciferol except H397Y. However, only H305F/H397Y and H305Y/H397F display maximal activation with cholecalciferol.

Table 4.1: EC_{50} values and Fold-activations with various ligands in HEK293T cells

Variant	Lithocholic Acid	Cholecalciferol	$1,25\ (OH)_2D_3$
wt	> 100 µM, 10 ± 5	-	3 nM, 82 ± 13
H305F	10 µM, 46 ± 22	10 µM, 39 ± 7	3 nM, 126 ± 43
H305Y	10 µM, 78 ± 15	10 µM, 70 ± 26	3 nM, 92 ± 16
H397F	-	100 µM, 24 ± 5	3 nM, 99 ± 23
H397Y	-	-	3 nM, 81 ± 15
H305F/H397F	3 ± 2	3 ± 2	3 ± 2
H305F/H397Y	10 µM, 50 ± 3	0.5 µM, 43 ± 5	3 nM, 46 ± 10
H305Y/H397F	10 µM, 47 ± 2	0.5 µM, 42 ± 16	3 nM, 52 ± 6
H305Y/H397Y	50 µM, 69 ± 9	50 µM, 60 ± 6	3 nM, 80 ± 15

- = no activation observed

All of the variants display activation with cholecalciferol except for H397F and H397Y. However, only H305F/H397Y and H305Y/H397F display maximal activation with cholecalciferol. In mammalian cell culture all of the variants were activated by cholecalciferol except H397Y and wild-type hVDR (Figure 4.3C). An EC_{50} of 500 nM was seen with H305F/H397Y and H305Y/H397F with 43 $\pm$5 and 42 $\pm$ 16 fold activations, respectively (Table 4.1). An EC_{50} of 10 µM was observed for H305F and H305Y as well as 39 $\pm$ 7 and 70 $\pm$ 26 fold activations, respectively. EC_{50}'s of 100 µM and 50 µM with fold activations of 24 $\pm$ 5 and 60 $\pm$ 6 were observed for H397F and H305Y/H397Y, respectively, and H305FH397F was constitutively active. Activation was not seen for wild-type hVDR. Maximal activation should only be observed with a combination of bulky residue and bulky residue with hydrogen bonding potential. Although activation was observed with H305F, H305Y, H305Y/H397Y, and H397F in mammalian cell culture, maximal activation was only observed with H305F/H397Y and H305Y/H397F (EC_{50} = 0.5 µM). These results confirm that a bulky residue in combination with a bulky residue with hydrogen bonding potential is needed at H305 and H397 in order to obtain maximal activation with cholecalciferol.

4.3 The Role of Volume and Hydrogen Bonding

Results have shown that a bulky residue is needed at H305 in order to achieve increased sensitivity towards LCA. Additional results have shown that a combination of bulky and bulky/hydrogen are needed in order to achieve maximal activation with cholecalciferol in yeast. In order to investigate the limits of volume and polarity that are necessary at H305 and H397 with LCA and cholecalciferol, another subset of single and double variants was created. In order to achieve this, residues were chosen that varied in

volume, hydrogen bonding potential, and charge. The residues chosen as substitutions were tryptophan, leucine, serine, and lysine. Tryptophan was chosen because this residue has the largest volume (227.8 $Å^3$) of all of the amino acids, testing the limit to which volume/bulk can be tolerated at those positions [6]. If volume/bulk is required at H305 for activation with LCA, then H305W should display ligand-activated growth with LCA, unless this mutation leads to constitutive activity. Also, because evidence suggests that the removal of hydrogen bonding potential at H305 does not affect activation with $1,25(OH)_2D_3$, ligand-activated growth should be observed with H305W. Leucine (166.7 $Å^3$) is slightly larger than histidine (153.2 $Å^3$) and therefore, the activation should vary with LCA, but should not affect activation with $1,25(OH)_2D_3$. Serine (89.0 $Å^3$) is smaller in volume than histidine but provides the potential of a hydrogen bond. This could lead to either an active or inactive receptor, depending on whether the volume/bulk or hydrogen bonding takes precedence. Lysine (168.6 $Å^3$) is a charged residue. Previously, results have shown that mutating H305 to a glutamine causes a decreased sensitivity with $1,25(OH)_2D_3$, leading to Type II Rickets [7]. Therefore, lysine variants should not display ligand-activated growth with $1,25(OH)_2D_3$, and is unlikely to be ligand-activated with LCA and cholecalciferol as well.

4.3.1 Testing Variants in Yeast

The variants were created via site-directed mutagenesis, then tested in yeast with $1,25(OH)_2D_3$, LCA, and cholecalciferol. As seen in Figure 4.4, only H305L shows ligand-activated growth with $1,25(OH)_2D_3$. An EC_{50} of 10 nM with a 14-fold activation, which is comparable to wild-type hVDR is observed. The lack of activation with H305W could be due to the bulkiness of this residue, indicating that perhaps a limitation

is present in terms of volume. The H305 single variants were also tested with LCA. As can be seen in Figure 4.4B, all of the variants showed ligand-activated growth comparable to wild-type hVDR, with an EC_{50} of 5 μM and a 3-fold activation, except in the case of the H305W and H305K variants. No ligand-activated growth is observed with H305K, which was expected based on previous results with another charged residue, H305Q. As can be seen in orange in Figure 4.4B, H305W displays an 8-fold activation with LCA, which demonstrates enhanced activity in comparison to wild-type hVDR and the other variants, which was expected due to the bulkiness of this residue. All of the variants were then tested with cholecalciferol and as can be observed in Figure 4.4C, no ligand-activated growth is occurs with any of the variants. These results were expected due to the fact that none of the single variants provided the combination of both bulk and bulk/hydrogen bonding potential at H305 and H397.

The H397 single variants were also tested in yeast with the same ligands (Figure 4.5). The variants were tested with $1,25(OH)_2D_3$ and ligand-activated growth was only observed with H397W and wild-type hVDR at and EC_{50} of 10 nM $1,25(OH)_2D_3$ with 11-fold activation (Figure 4.5A). The variants were then tested with LCA and all of the variants showed similar EC_{50}'s to wild-type hVDR of 5 μM LCA and 3-fold activation, which was expected due to the hypothesis is H305 is the most important residue for LCA ligand activation (Figure 4.5B). When the variants were tested with cholecalciferol, none displayed ligand-activated growth, as was expected (Figure 4.5C).

The H305/H397 double variants were also tested in yeast (Figure 4.6). The variants were tested with $1,25(OH)_2D_3$, and as can be seen in Figure 4.6A, only H305W/H397W

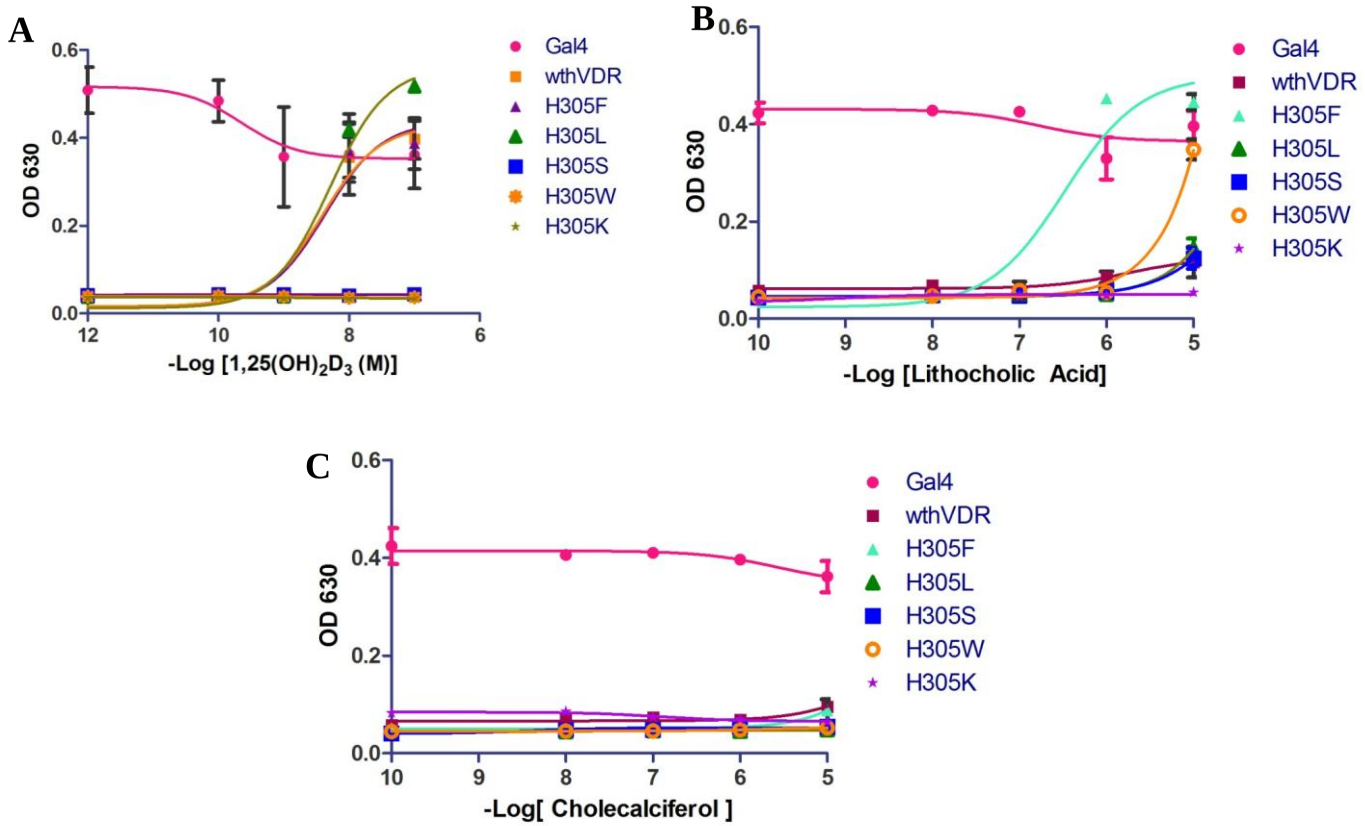

Figure 4.4: Testing H305 single variants with various ligands in yeast. A)Ligand-activated growth was observed with wild-type hVDR and H305L with 1,25$(OH)_2D_3$ in adenine selective media. B) Enhanced ligand-activated growth is observed for H305W in comparison to wild-type hVDR with lithocholic acid in histidine selective media with 0.1 mM 3AT. C) No ligand-activated growth is observed with cholecalciferol in histidine selective media with 0.1 mM 3AT.

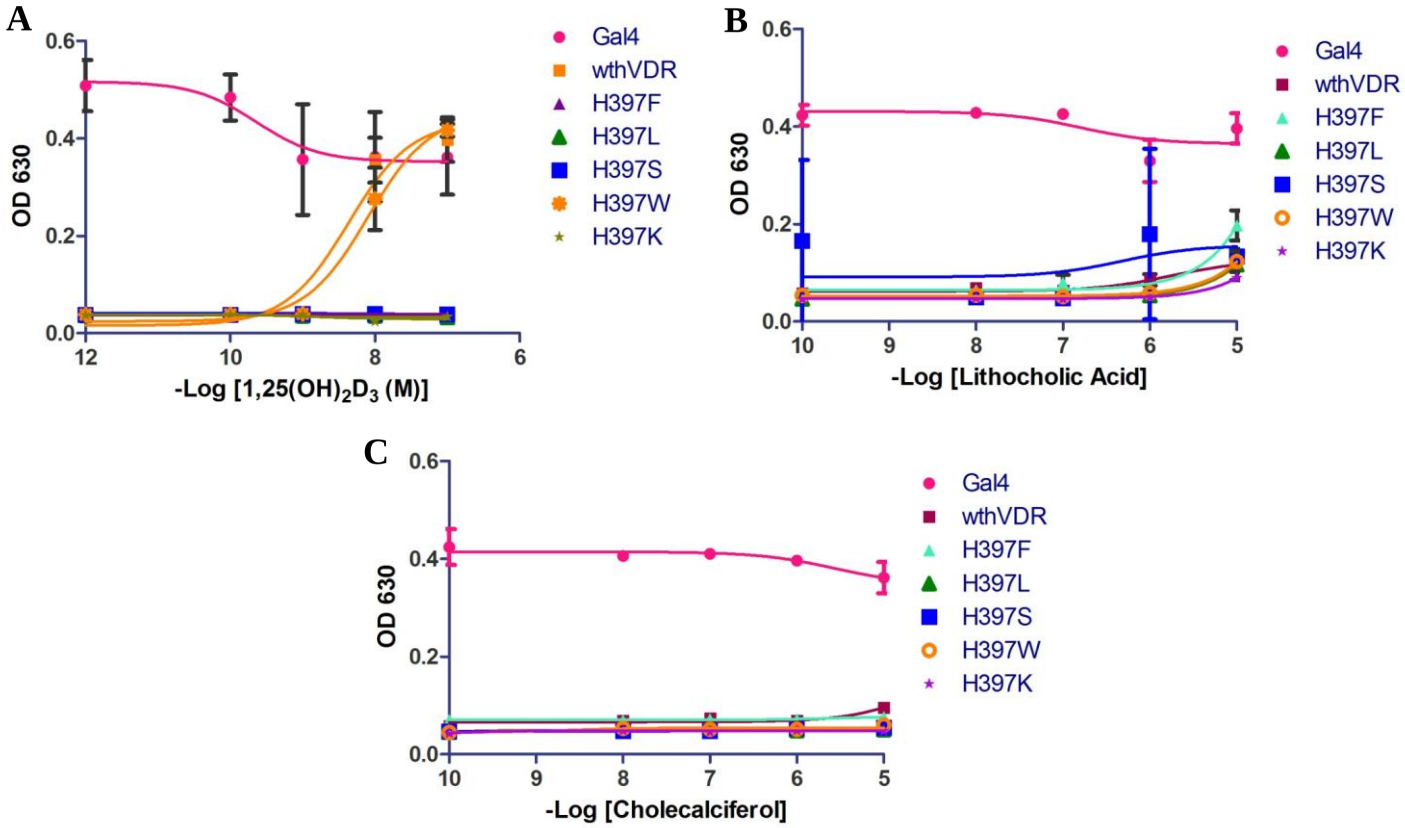

Figure 4.5: Testing H397 single variants with various ligands in yeast. A) Ligand-activated growth is observed for wild-type hVDR and H397W with 1,25$(OH)_2D_3$ in adenine selective media. B) All variants display wild-type like growth with lithocholic acid in histidine selective media with 0.1 mM 3AT. C) No ligand-activated growth is observed with cholecalciferol in histidine selective media with 0.1 mM 3AT.

shows ligand-activated growth. H305W/H397W displayed an EC_{50} of 50 nM and 6-fold activation with $1,25(OH)_2D_3$ in comparison to an EC_{50} of 10 nM and 11-fold activation with wild-type hVDR. Therefore, a ten-fold decrease is observed with this double variant with $1,25(OH)_2D_3$. This was not expected because H305F/H397F is constitutively active, and H305W/H397W was hypothesized to also be constitutively active due to the fact that tryptophan is bulkier than phenylalanine should prevent the ligand from gaining access to the pocket. When other double variants were tested with LCA, none showed ligand-activated growth (Figure 4.6B). H305W/H397W seems to display slight growth with cholecalciferol with an EC_{50} of 5 µM and 3-fold activation, most likely due to the bulkiness of the two residues. There is a possibility that because tryptophan is so bulky, the introduction of two tryptophans creates novel interactions with cholecalciferol that compensates for the lack of hydrogen bonding potential.

4.3.2 Testing Variants in Mammalian Cells

Since all of the H305 singles and H305W/H397W showed some growth with either LCA or cholecalciferol, each variant was tested in mammalian cell culture. When the single variants were tested at higher concentrations of $1,25(OH)_2D_3$, activation was observed with all of the variants. Wild-type hVDR, H305F, and H305L all display EC_{50} of 3 nM, whereas H305S, H305W, and H305K show EC_{50} of 100 nM with $1,25(OH)_2D_3$ with fold activations ranging from 126 ± 43 for H305F to 44 ± 9 for H305W confirming that hydrogen bonding is not the predominant factor in ligand activation with $1,25(OH)_2D_3$ (Figure 4.7A, Table 4.2). H397 also appears to be less tolerant to mutations than H305. However, H305K (charged), H305S (least bulky), and H305W (most bulky)

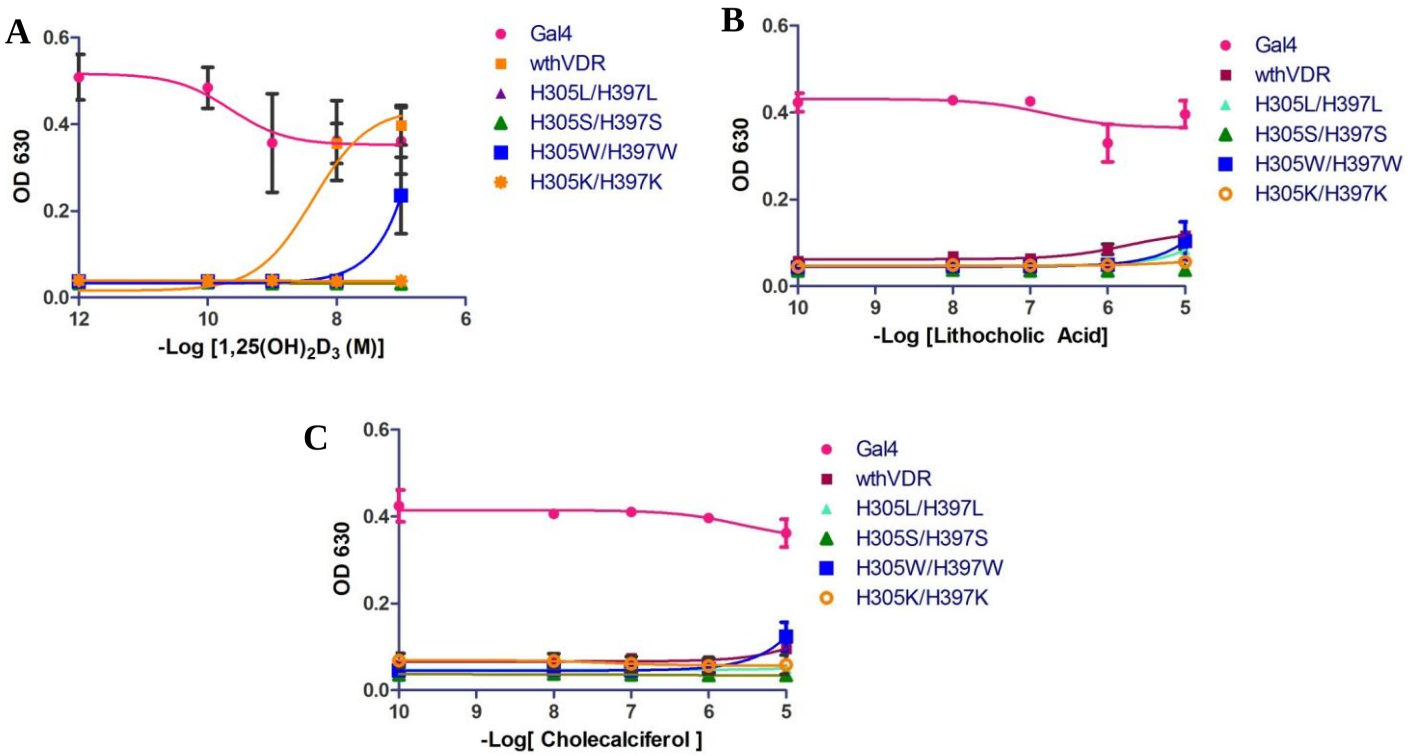

Figure 4.6: Testing H305 and H397 double variants with various ligands in yeast. A) H305W/H397W displays a ten-fold decreased sensitivity in comparison to wild-type with $1,25(OH)_2D_3$ in adenine selective media. B) None of the variants display ligand-activated growth with lithocholic acid in histidine selective media with 0.1 mM 3AT. C) No ligand-activated growth is observed with cholecalciferol in histidine selective media with 0.1 mM 3AT.

show a 10-fold decrease in sensitivity with $1,25(OH)_2D_3$ suggesting that hVDR prefers residues that have the hydrogen bonding potential (histidine or tyrosine) or residues that are similar or slightly larger in volume than histidine (leucine and phenylalanine). Charged residues, residues that are significantly smaller in volume than histidine, and extremely bulky residues do not appear to tolerate $1,25(OH)_2D_3$ as the ligand.

When the single variants were tested with LCA, activation was observed with H305W and wild-type hVDR whereas the other single variants were not ligand-activated (Figure 4.7B). H305W has an EC_{50} of 30 μM and 14 $\pm$ 22 fold activation with LCA in comparison to 100 μM and 8 $\pm$ 1 fold activation with wild-type hVDR, which appears to confirm the yeast data (Table 4.2). H305W was the only variant that displayed ligand activation with LCA, confirming our hypothesis that bulkiness is required at H305 to obtain activation with LCA. However, H305W has a higher EC_{50} with LCA than H305F and H305Y suggesting that these residues may be the volume at which maximal activation occurs; therefore, a residue bulkier than histidine is required for activation with LCA. However, if the residue is too bulky the result is most likely crowding, which could lead to a decrease in ligand sensitivity.

When the variants were tested with cholecalciferol activation was observed with H305L displaying an EC_{50} of 30 μM and 29 $\pm$ 5 fold activation, which was not observed in chemical complementation. No ligand-activated growth was observed with cholecalciferol and H305L in yeast. This inconsistency may be a result of the coactivator differences in the two systems. A specific coactivator is introduced into yeast, whereas, mammalian cells have various coactivators present. H305L shows the same trend as H305F and H305Y suggesting that H305L, while less bulky that histidine continues to

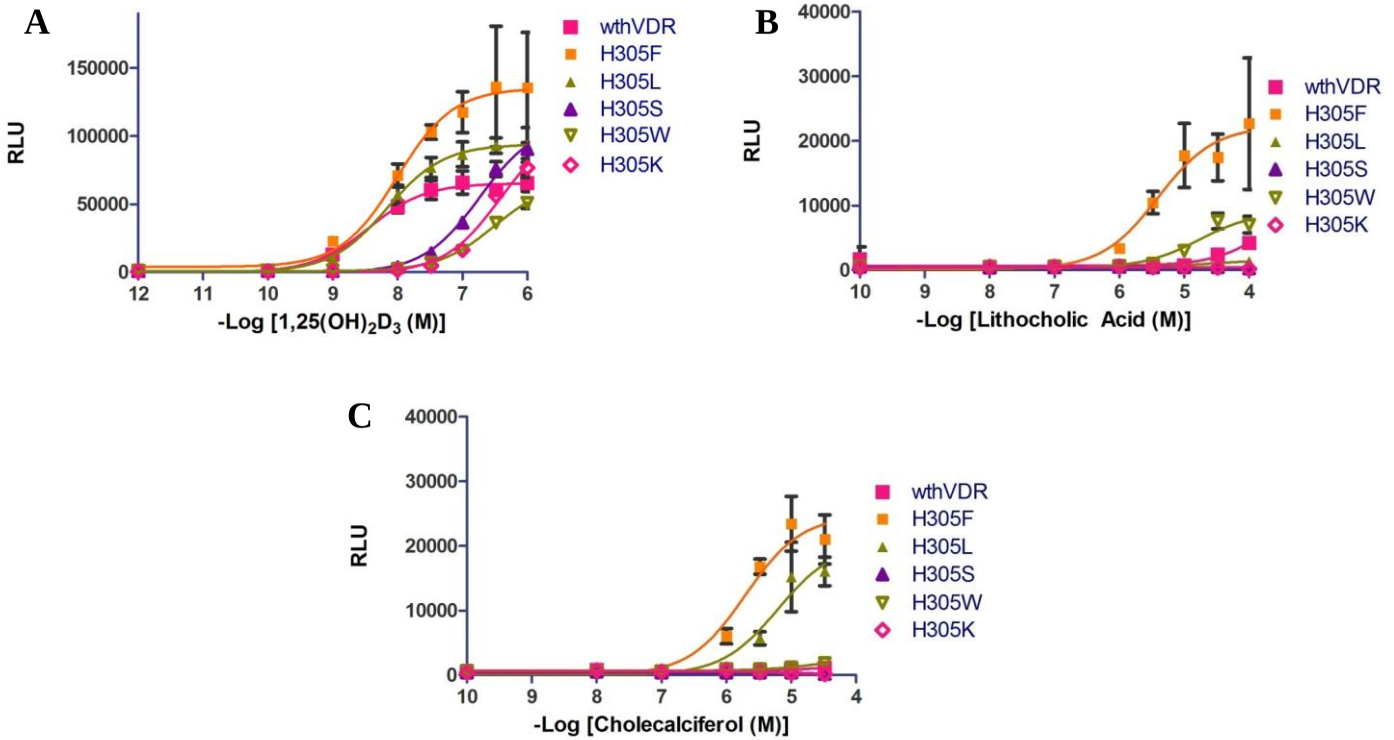

Figure 4.7: Testing H305 single variants with various ligands in mammalian cells. A) All of the variants display activation with $1,25(OH)_2D_3$. B) Only H305W displays activation with lithocholic acid. C) H305L displays activation with cholecalciferol.

Table 4.2: EC_{50} values and Fold-activations with various ligands in HEK293T cells

Variant	Lithocholic Acid	Cholecalciferol	$1,25\ (OH)_2D_3$
wt	> 100 µM, 8 ± 1	-	3 nM, 82 ± 13
H305F	10 µM, 46 ± 22	10 µM, 39 ± 7	3 nM, 126 ± 43
H305L	-	30 µM, 29 ± 5	3 nM, 89 ± 16
H305S	-	-	100 nM, 103 ± 20
H305W	30 µM, 14 ± 5	-	100 nM, 44 ± 9
H305K	-	-	100 nM, 104 ± 10
H305W/H397W	-	3 µM, 21 ± 6	100 nM, 51 ± 15
H305W/H397F	3 µM, 11 ± 1	1 µM, 15 ± 2	100 nM, 27 ± 5
H305L/H397Y	-	1 µM, 23 ± 2	10 nM, 55 ± 7

- = no activation observed

make other LBP contacts crucial for activation with cholecalciferol. Activation was also observed with H305W/H397W at an EC_{50} of 3 μM and 21 $\pm$ 6 fold activation with cholecalciferol. These results suggest that bulkiness plays a more critical role in ligand activation at H305 and H397 than hydrogen bonding and electrostatics. The hypothesis was that H305W/H397W would most likely be constitutively active since H305F/H397F displays constitutive activity, and tryptophan is a bulkier residue than phenylalanine. However, H305W/H397W displays ligand-activated growth, suggesting that the interactions of the two tryptophans with the surrounding residues are different that the interactions with the two phenylalanines, leading to the different activation profiles.

4.4 Mixed Double Variants

Based on the results mentioned above, double variants were created to further test the hypotheses that were made. By combining mutations, a more comprehensive understanding of the residues that are necessary at H305 and H397 to achieve activation with LCA and cholecalciferol could be achieved. Therefore, the tryptophan, leucine, serine, tyrosine, and phenylalanine residues were tested in combination with each other at H305 and H397.

The variants were created via site-directed mutagenesis and tested in liquid quantitation in yeast. Figure 4.8 shows a subset of the variants that were created. As can be seen in Figure 4.8A, H305W/H397F displays an EC_{50} of 0.5 μM and 6-fold activation with LCA in comparison to an EC_{50} of 5 μM and 3-fold activation with wild-type hVDR and H305F/H397L. Only two of the variants displayed ligand-activated growth with cholecalciferol; H305W/H397F and H305L/H397Y showed EC_{50}'s of 1 μM and 5 μM, respectively and 5-fold activation with cholecalciferol (Figure 4.8B).

Based on these results, H305W/H397F and H305L/H397Y were tested in mammalian cell culture with 1,25$(OH)_2D_3$, LCA, and cholecalciferol. As can be seen in Figure 4.9A, both variants display a decreased sensitivity towards 1,25$(OH)_2D_3$. H305W/H397F and H305L/H397Y display EC_{50}'s of 100 nM and 10 nM with fold activations of 27 $\pm$ 5 and 55 $\pm$ 7, respectively. The decrease in activation with H305W/H397F was expected due to previous results showing a reduced sensitivity with 1,25$(OH)_2D_3$ in regards to the H305W and H397F single variants. However, the decreased sensitivity with H305L/H397Y was not expected due to the previous results that showed wild-type activation levels with both of the single variants, H305L and H397Y. The possibility exists that the combination of the two residues leads to changes in the contacts with other residues that affect the sensitivity.

As observed in yeast, H305W/H397F displays activation with LCA with an EC_{50} of 3 µM and 11 $\pm$1 fold activation in comparison to 100 µM and 8 $\pm$ 1 fold activation for wild-type hVDR in mammalian cells (Figure 4.9B, Table 4.2). These results were expected due to previous results that showed activation of H305W with LCA. No ligand activation was observed with H305L/H397Y, which was expected due to the lack of activation of both single variants with LCA.

Both H305W/H397F and H305L/H397Y display activation with cholecalciferol with an EC_{50} of 1µM and 15 $\pm$ 2 and 23 $\pm$ 2 fold activations, respectively as can be seen in Figure 4.8C (Table 4.2). These results were expected due to the fact that both the H305L single variant and the H305W/H397W double variant display activation with cholecalciferol. Therefore, H305L in combination with H397Y displaying increased sensitivity over the single variants alone is not surprising. Combining H305W with

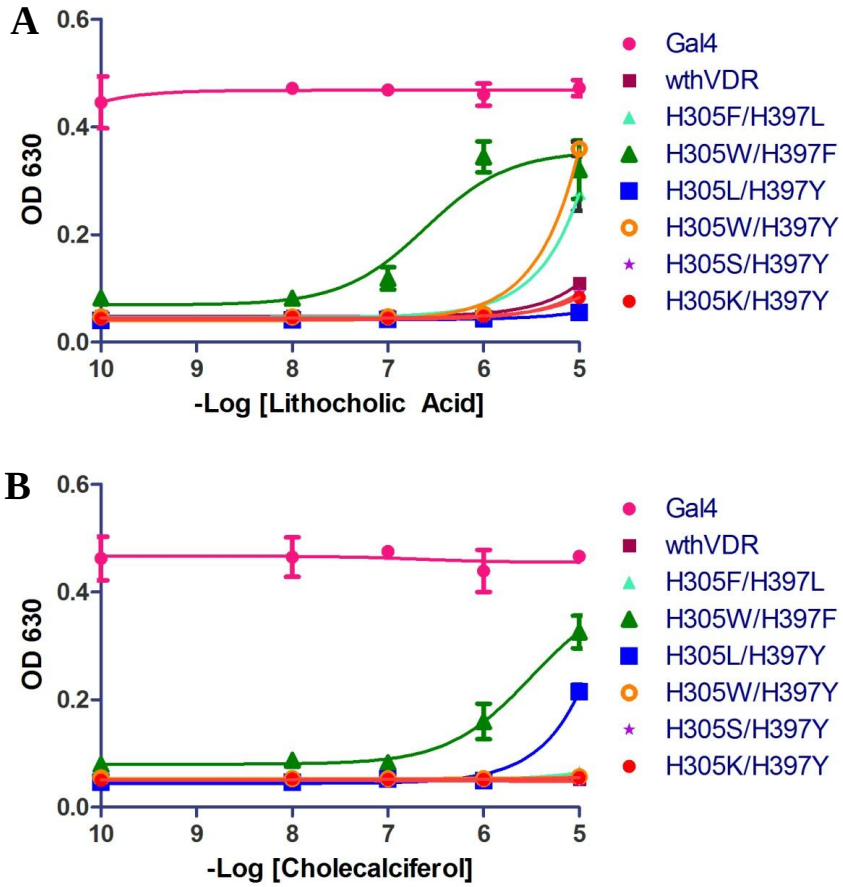

Figure 4.8: Testing H305 and H397 mixed double variants with various ligands in yeast. A) H305W/H397Y, H305F/H397L, and H305W/H397F display ligand-activated growth with lithocholic acid in histidine selective media with 0.1 mM 3AT. B) H305W/H397F and H305L/H397Y display ligand-activated growth with cholecalciferol in histidine selective media with 0.1 mM 3AT.

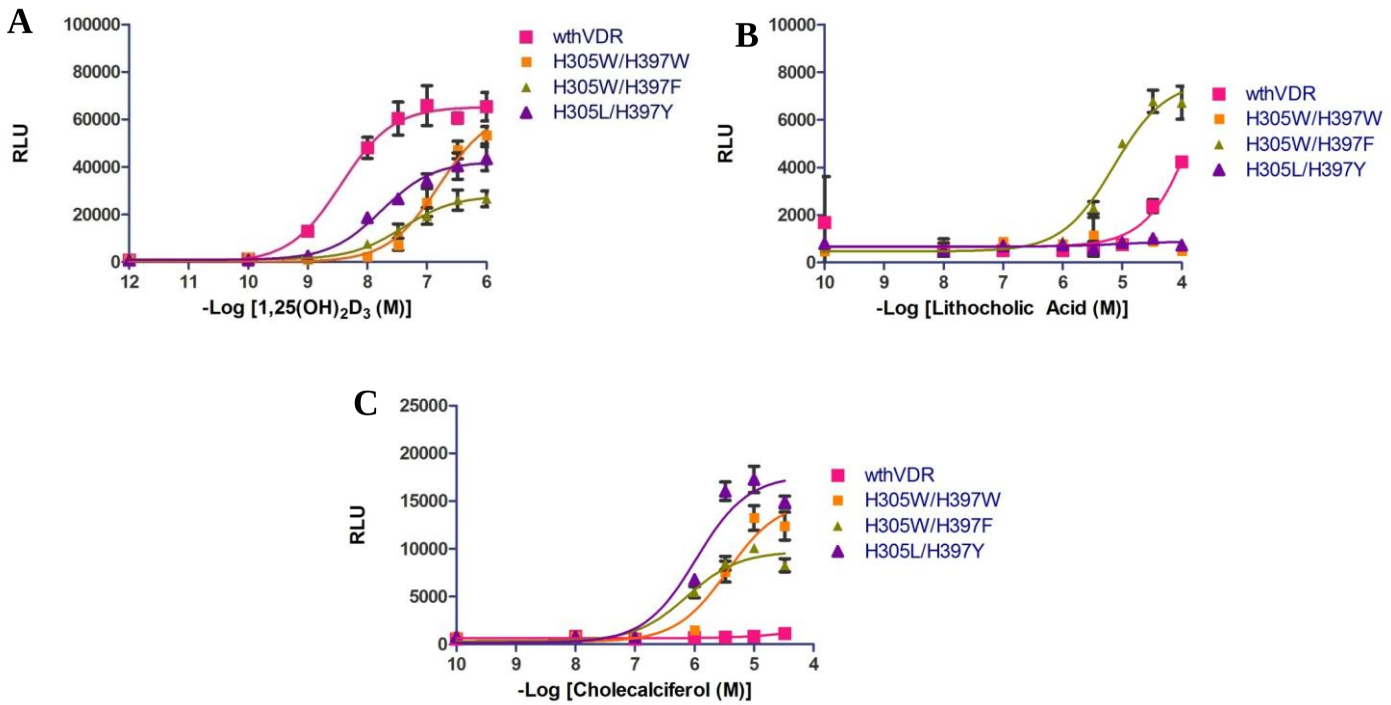

Figure 4.9: Testing Various H305 and H397 double variants with various ligands in mammalian cells. A) All variants display decreased activation in comparison to wild-type with $1,25(OH)_2D_3$. B) H305W/H397F displays increased sensitivity towards lithocholic acid. C) All variants display activation with cholecalciferol.

H397F maintains bulkiness at both positions and achieves similar activation to the H305W/H397W variant with cholecalciferol. These results confirm that bulkiness is more important than hydrogen bonding and electrostatics at these two positions.

4.5 Summary

The results of all of the H305 and H397 variants suggest that certain trends can be observed with each ligand. The results suggest that hydrogen bonding is not the most important component of ligand activation with $1,25(OH)_2D_3$. However, H397 seems to be less tolerable to mutations than H305 when binding to $1,25(OH)_2D_3$, but hydrophobic residues are tolerated. Activation by LCA appears to require bulkiness at H305 due to the fact that each variant that is activated by LCA at lower concentrations than wild-type hVDR contains a residue bulkier than histidine at 305. Combining all of the results with cholecalciferol, maximal activation is still observed with a combination of bulky and bulky/hydrogen bonding as is observed with H305F/H397Y and H305Y/H397F. However, in order to obtain the slightest activation with cholecalciferol, bulky residues , such as phenylalanine and tyrosine are needed at both 305 and 397, and specifically at H397. These results suggest that H305 and H397 are important residues in ligand binding and activation, given that certain combinations of residues at these positions are necessary for activation with the different ligands.

4.6 Methods and Materials

4.6.1 Ligands

Lithocholic acid, cholecalciferol, and 1α,25-dihydroxyvitamin D_3 were ordered from MP Biomedicals, LLC (Solon, OH), Sigma-Aldrich (St. Louis, MO), and BIOMOL (Plymouth Meeting, PA), respectively. 10 mM stocks of lithocholic acid and

cholecalciferol and a 13.3 μM stock of 1α,25-dihydroxyvitamin D_3 were made with 80% ethanol: 20% DMSO as the solvent and stored at 4 °C.

4.6.2 Site-directed Mutagenesis

pGBDhVDR was used as the template DNA, site-directed mutagenesis with primers: 5'-gcgcagcctcaatgaggagCACtccaagcagtatcgatgcctctccttccagcctg-3' and 5'-caggctggaaggagaggcatcgatactgcttggaGTGctcctcattgaggctgcgc-3' for the H397 variants and 5'-caagtaccgcgtcagtgacgtcaccaaagccggaCACagcctggagctgattgagcc-3' and 5'-ggctcaatcagctccaggctGTGtccggctttggtgacgtcactgacgcggtacttg-3' for the H305 variants. The CAC and GTG were changed to codons for the respective variant. The PCR program used was: 95 ºC for 4 min, 95 ºC for 1 min, 56 ºC for 1 min, 72 ºC for 9 min, and 18 cycles.

4.6.3 Liquid Quantitation Assay

Variants were tested in liquid quantitation assays in 96-well plates with media lacking adenine, leucine, and tryptophan (SC-ALW), with or without LCA, cholecalciferol, or $1,25(OH)_2D_3$ at varying concentrations (ranging from 0.01 μM - 10 μM for LCA and cholecalciferol, 1 nM - 1 μM for $1,25(OH)_2D_3$). A 4:1 ratio of media (SC-ALW): cells (yeast resuspended in water) were aliquoted into 96-well plates. Plates were incubated at 30°C, with shaking at 170 rpm. Optical density (OD) readings at 630 nm were recorded at 0, 24, and 48 hours as a measure of growth density.

4.6.4 Mammalian Cell Culture

HEK293T cells (ATCC, USA) were transfected with the plasmids of interest. These plasmids contain the Gal4DBD (GBD) fused to the corresponding VDR ligand binding domain (GBD:LBD fusion under the control of a cytomegalovirus (CMV) promoter).

The reporter plasmids were p17*4TATAluc, containing the *Renilla* luciferase gene under the control of four Gal4 response elements located upstream from a minimal thymidine kinase promoter, and pCMXβgal, a plasmid containing the β-galactosidase gene under the control of the mammalian CMV promoter. Lipofectamine 2000 (Invitrogen, USA) served as the cationic lipid and transfection experimental details are described in Taylor *et al*. [8]. The ligands were added to the wells at various concentrations ((0.01 μM- 100 μM) LCA and (0.01 μM- 32 μM) cholecalciferol). Cells were harvested and analyzed for luciferase and β-galactosidase activity. All data points represent the average of triplicate experiments normalized against β-galactosidase activity. Error bars represent the standard deviation calculated using standard deviation: σ = Square root ($\Sigma[(X-\mu)^2]/N$). Fold activation was calculated by dividing the value at maximal activation by the value at the no ligand data point.

4.7 References

1. Trott O, Olson AJ: **AutoDock Vina: Improving the speed and accuracy of docking with a new scoring function, efficient optimization, and multithreading**. *J Comput Chem* 2009:456-461.

2. Rochel N, Wurtz JM, Mitschler A, Klaholz B, Moras D: **The crystal structure of the nuclear receptor for vitamin D bound to its natural ligand**. *Mol Cell* 2000, **5**(1):173-179.

3. Mizwicki MT, Bula CM, Bishop JE, Norman AW: **New insights into Vitamin D sterol-VDR proteolysis, allostery, structure-function from the perspective of a conformational ensemble model**. *J Steroid Biochem Mol Biol* 2007, **103**(3-5):243-262.

4. Mizwicki MT, Bula CM, Mahinthichaichan P, Henry HL, Ishizuka S, Norman AW: **On the mechanism underlying (23*S*)-25-dehydro-1alpha(OH)-vitamin D3-26,23-lactone antagonism of hVDRwt gene activation and its switch to a superagonist**. *J Biol Chem* 2009, **284**(52):36292-36301.

5. Jones G, Strugnell SA, DeLuca HF: **Current understanding of the molecular actions of vitamin D**. *Physiol Rev* 1998, **78**(4):1193-1231.

6. Reichert J, Suhnel J: **The IMB Jena image library of biological macromolecules: 2002 update**. *Nucleic Acids Res* 2002, **30**(1):253-254.

7. Rochel N, Hourai S, Moras D: **Crystal structure of hereditary vitamin D-resistant rickets-Associated mutant H305Q of vitamin D nuclear receptor bound to its natural ligand**. *J Steroid Biochem Mol Biol*.

8. Taylor JL, Rohatgi P, Spencer HT, Doyle DF, Azizi B: **Characterization of a molecular switch system that regulates gene expression in mammalian cells through a small molecule**. *BMC Biotechnol*, **10**:12.

CHAPTER 5
EXPRESSION AND PURIFICATION OF HVDR AND HVDR VARIANTS

5.1 Introduction

In order to further study wild-type hVDR and the hVDR variants, they needed to be expressed and purified in bacteria using a maltose binding protein fusion. The maltose binding protein is used throughout protein expression to allow for the protein of interest to be expressed in the soluble fraction of protein within the bacteria. Although the exact mechanism by which the maltose binding protein aids in solubility is unknown, the protein is hypothesized to act as a chaperone in order to move the protein of interest into the soluble fraction [1].

Once wild-type hVDR and the hVDR variants are expressed and purified various studies can be conducted on the pure proteins. Ligand binding studies can be performed to determine the binding affinity of various ligands. Also, an attempt can be made to crystallize the variants with the various liagnds of interest to allow for a more comprehensive structural analysis of the protein-ligand interactions.

5.2 Expression and Purification of Proteins

5.2.1 Cloning VDR Ligand Binding Domains into a Bacterial Expression Vector

In order to perform binding studies on wild-type hVDR and the variants discussed in Chapters 3 and 4, the proteins were cloned into pMalRXR, a bacterial expression vector. The pMal vector contains the gene for the maltose binding protein (MBP) fused to the RXRLBD. The MBP increases the solubility of proteins expressed in *E.coli*, which is necessary for protein purification. The mechanism by which MBP aids in solubility is

not well understood, but several hypotheses indicate that the MBP could act as a chaperone as well as engage in preventing aggregation [1].

The wild-type hVDR ligand binding domain (LBD) and variants: H305FLBD, H305F/H397YLBD, and H305FH397FLBD were cloned out of pGBDhVDR with primers containing *EcoR*I and *Spe*I restriction sites. pMalRXR was then digested with *EcoR*I and *Spe*I , which removed the RXR gene and ligated wild-type hVDR and the variants into the vector. These plasmids were confirmed by sequencing.

5.2.2 Expression and Purification of MBP Fusion and TEV Protease

Once the bacterial expression vectors were created, the proteins were over-expressed and purified using an amylose resin column. MBP generally binds to maltose, but can also bind to amylose [2]. Therefore, when the over expressed protein is applied to the resin, the fusion protein will bind to the amylose resin. The fusion protein consists of the MBP fused to the VDRLBD. The two proteins are separated by a histidine-tag and a TEV protease cleavage site (Figure 5.1).

Throughout the expression and purification process, samples from each step are collected for further use. These samples are run on a 12 % SDS-PAGE gel to determine the success of protein expression and purification. As can be seen in Figure 5.2, lanes 2 and 3 represent the bacterial expression. A bright band is observed at approximately 80 kD in lane 2, which represents the over expression of the fusion protein. The insoluble fraction in lane 3 also indicates the presence of the fusion protein, which is most likely due to incomplete cell lysis during the expression. Lanes 4-8 represent the purification, where lane 4 represents the column flow through of the soluble fraction. The fusion

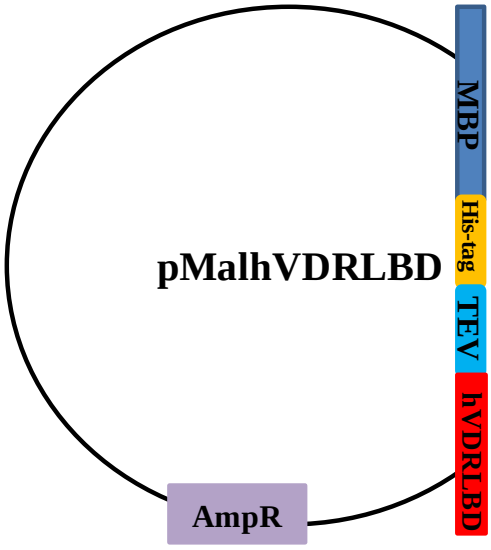

Figure 5.1: Diagram of fusion protein plasmid.

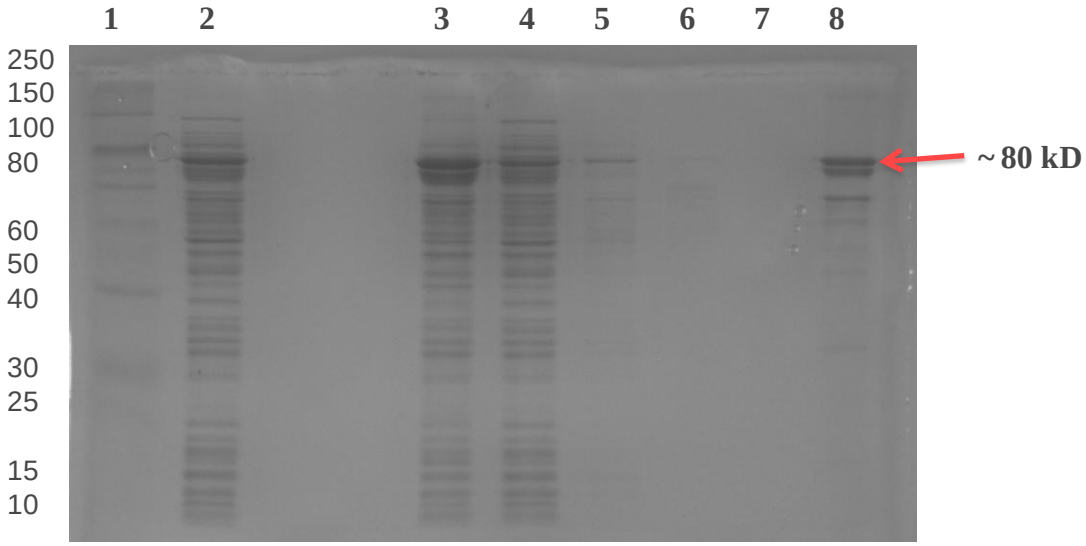

Figure 5.2: Expression and Purification of MBP-VDRLBD fusion protein. The fusion protein is about 80 kD. Lane 1= Protein marker, Lane 2= soluble fraction, Lane 3= insoluble fraction, Lane 4= post resin, Lanes 5-7= resin washes, Lane 8 = purified fusion protein.

protein of the MBP and VDR is observed in the post resin, which indicates unbound protein within the resin, implying more resin might be needed to accommodate the amount of protein present in the soluble fraction. The purified fusion protein is observed at approximately 80 kD (lane 8), yielding 2-3 mg/mL of purified fusion protein. Binding studies require pure VDRLBD; the MBP must be cleaved from the VDR.

The tobacco etch virus (TEV) protease functions like a serine protease, with the exception that serine in the active site replaced by a cysteine in TEV. The TEV protease cleaves via peptide hydrolysis at a specific sequence, EXXYXQ(G/S), where cleavage occurs between the glutamate and glycine or serine [3]. TEV protease was expressed and purified using a metal affinity column, due to the histidine-tag present at the N-terminal of this protein. The metal affinity column has space requirements that only allow proteins with adjacent histidines to bind to the resin. Expressed and purified TEV protease should display a band around 27 kD. As can be seen in Figure 5.3, TEV protease can be observed in the soluble fraction in lane 2 at around 27 kD. However, residual protease is present in the insoluble fraction (lane 3) in comparison to the soluble fraction, indicating incomplete cell lysis. Purified TEV protease is seen at 27 kD (lane 7) with a concentration of approximately 1 mg/mL.

5.2.3 Fusion Protein Cleavage and Purification

Once the TEV protease was purified, the protease was used to cleave the fusion of MBP and the wild-type VDR or the VDR variants. The TEV protease was added to the fusion protein at a ratio of 1 mg of TEV protease for every 100 mg of the fusion protein and incubated for 24 hours at 4 °C. TEV protease is most active at room temperature,

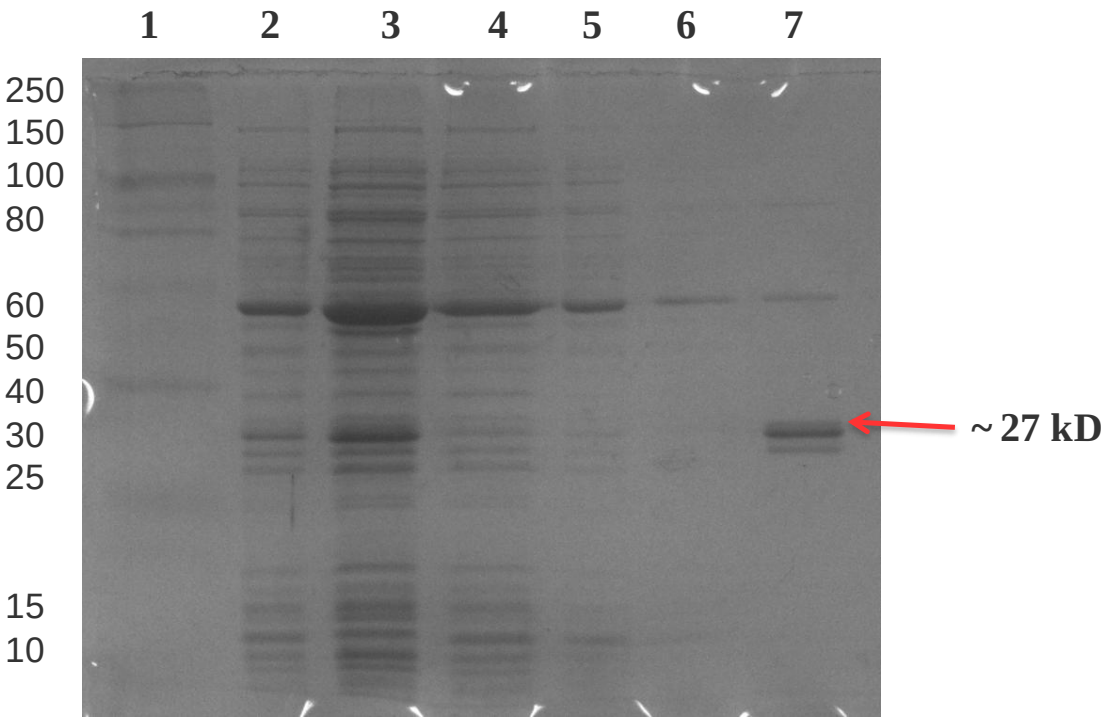

Figure 5.3: Expression and Purification of TEV protease. TEV protease should be about 27 kD. Lane 1= protein marker, Lane 2= soluble fraction, Lane 3= insoluble fraction, Lane 4= post resin, Lanes 5 and 6= resin washes, Lane 7 = purified TEV protease

however cleavage was performed at 4 °C to maintain the receptor's stability, and the time for cleavage was extended to 24 hours. After the 24 hour digest, the cleaved sample was run on a metal affinity column to purify the VDRLBD's, eliminating any remnants of the MBP and TEV, since both also contain His-tags.

As can be observed in Figure 5.4, the band at 80 kD observed in lane 2 shows that the protein was cleaved but that the cleavage was incomplete, indicating the presence of fusion protein. Lane 3 shows the flow-through of the cleavage, with pure VDRLBD at 35 kD along with residual fusion still present at 80 kD. Lane 4 shows that MBP is present at 37 kD in the column elution. Lanes 5-7 show the same results but with the H305F variant. The purified LBDs fluctuated in concentration from 0.8 to1.5 mg/mL after cleavage and purification. Wild-type hVDR and variants have been expressed and purified successfully, but contain remnants of the fusion protein, preventing a higher yield of the pure protein.

5.3 Summary and Future Work

Wild-type hVDR and hVDR variant proteins were able to be obtained via expression and purification. However, pure protein could not be isolated due to the presence of fusion protein in the sample after the cleavage occurred. Therefore, optimization of the TEV protease expression and purification is needed to ensure consistent activity of this protease. Also, the cleavage time and ratio of TEV protease to fusion protein need to be further investigated. Once pure protein can be isolated, further studies can be performed with wild-type hVDR and the hVDR variants.

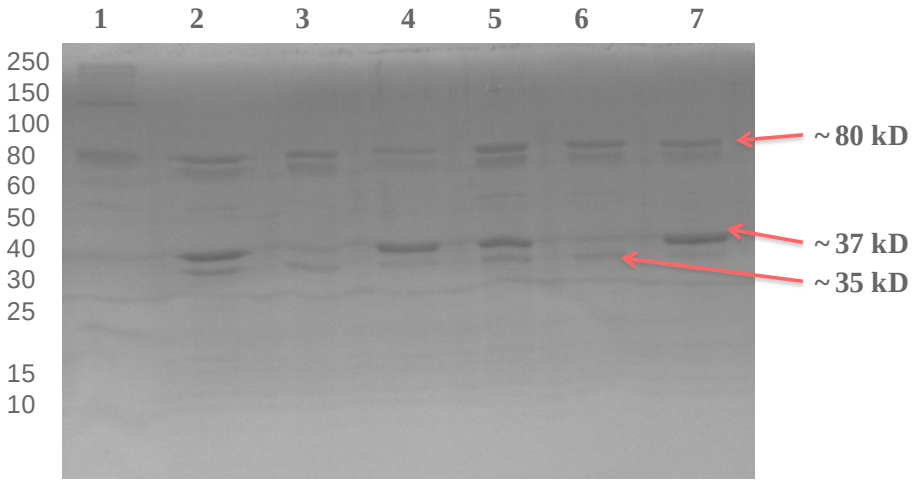

Figure 5.4: Purification of Cleaved Fusion Protein. The fusion protein is about 80 kD. However, the Maltose binding protein and the VDRLBD are about 37 and 35 kD, respectively. Lane 1= protein marker, Lane 2= wthVDR cleaved, Lane 3= pure VDRLBD, Lane 4= MBP elution, and Lanes 5-7 correspond to lanes 2-4 but with H305FLBD.

5.4 Methods and Materials

5.4.1 Ligands

Lithocholic acid, cholecalciferol, and 1α,25-dihydroxyvitamin D_3 were ordered from MP Biomedicals, LLC (Solon, OH), Sigma (St. Louis, MO), and BIOMOL (Plymouth Meeting, PA), respectively. 10 mM stocks of lithocholic acid and cholecalciferol and a 13.3 μM stock of 1α,25-dihydroxyvitamin D_3 were made with 80% ethanol: 20% DMSO and stored at 4 °C.

5.4.2 Cloning VDRLBD's into pMalRXR

The wild-type hVDRLBD and hVDR variant LBD's were amplified with PCR using the following primers: 5'- ccg gaattc gacagtctgcggcccacgc-3' and 5'- gg actagt tcaggagatctcattgccaaaca-3'. The underlined sequences denote *EcoR*I and *Spe*I restriction sites, respectively. pMalRXR was digested with the two restriction enzymes, removing the RXRLBD from the plasmid. The amplified VDRLBD's were then ligated into the linearized pMal plasmid, creating pMalwthVDRLBD, pMalH305FLBD, pMalH305F/H397YLBD, and pMalH305F/H397FLBD. The plasmids were confirmed by sequencing (Operon, USA).

5.4.3 Expression and Purification of MBP-VDR LBD Fusions

The VDRLBD's were overexpressed in *E.coli* BL21D3 cells. The cells were grown to an OD of approximately 0.60 at room temperature and 300 rpm, then induced with 0.300 mM IPTG. The induction was performed at room temperature for 16 hours, while shaking at 300 rpm. The overnight cultures were pelleted and lysed to collect the soluble fraction for purification. The purification was performed with an amylose resin affinity

column (New England Biolabs, MA), and dialysed overnight. The protein was then run on a 12 % SDS-Page gel and protein concentrations were determined by Bradford assay.

The maltose wash buffer was prepared as follows: 200 mM NaCl, 10 mM Na_2PO_4, 10 mM K_2PO_4, 1 mM EDTA, and 0.1 mM maltose. The maltose elution buffer was prepared in the same manner as the wash buffer with the exception of the maltose concentration. The elution buffer had a maltose concentration of 20 mM. The dialysis buffer was prepared as follows: 150 mM NaCl, 80 mM Tris-HCl pH 7.6, and 35 % glycerol.

5.4.4 TEV Protease Expression and Purification

TEV protease was overexpressed in *E.coli* BL21-RIL D3 cells. The cells were grown to an OD of about 0.60 at 37 °C and 300 rpm and then induced with 1 mM IPTG. The induction was performed for 4 hours at 30 °C and 300 rpm. After induction the cells were pelleted and lysed and the soluble fraction used for purification. TEV was purified using a metal affinity column. The purified protein was run on a 12% SDS-PAGE gel and the concentration was determined by Bradford assay.

The TEV protease wash buffer was prepared as follows: 50 mM Na_2PO_4, 50 mM K_2PO_4, 100 mM NaCl, 35 % glycerol, and 25 mM imidazole. The TEV protease elution buffer was prepared the same way as the wash buffer except for a differing concentration of imidazole. The imidazole concentration of the elution buffer was 500 mM.

5.4.5 Cleaving and Purifying the Fusion Protein

TEV protease was added to the dialysed fusion protein in a ratio of 1 mg of TEV protease for every 100 mg of fusion protein. The sample was then mixed for 24 hours at 4 °C. The sample was then run through a metal affinity column and the flow through was

collected (pure VDRLBD). The column was washed with TEV protease elution buffer to collect the cleaved MBP. The protein was then run on a 12 % SDS-PAGE gel and the concentration determined by Bradford assay.

5.5 References

1. Fox JD, Waugh DS: **Maltose-binding protein as a solubility enhancer**. In: *Methods in Molecular Biology*. Clifton: Springer Press; 2003.

2. Maina CV, Riggs PD, Grandea AG, 3rd, Slatko BE, Moran LS, Tagliamonte JA, McReynolds LA, Guan CD: **An *Escherichia coli* vector to express and purify foreign proteins by fusion to and separation from maltose-binding protein**. *Gene* 1988, **74**(2):365-373.

3. Phan J, Zdanov A, Evdokimov AG, Tropea JE, Peters HK, 3rd, Kapust RB, Li M, Wlodawer A, Waugh DS: **Structural basis for the substrate specificity of tobacco etch virus protease**. *J Biol Chem* 2002, **277**(52):50564-50572.

CHAPTER 6
Applications of Chemical Complementation: Enzyme-Activated Growth

6.1 Introduction

Chemical complementation is a genetic selection system which links the survival of yeast to the function of a nuclear receptor and a small molecule [1-4]. This versatile system can be used for a number of applications, ranging from drug discovery to protein and enzyme engineering [2, 3]. The ability to employ a genetic selection system to evaluate a library of engineered enzymes for novel activity is an extremely useful tool. This method could facilitate the creation of novel small molecules biosynthetically, as an alternative to synthetic methods. For example, a number of antibiotics, such as oxacillin and cloxacillin, are currently limited to synthetic production, potentially leading to environmental and economic concerns, such as waste production. Thus, chemical complementation is a tool for biocatalysis, where the product of catalysis is linked to the survival of yeast, hence enzyme-activated growth [5-7].

In order to use chemical complementation for enzyme engineering and enzyme-activated growth, several components must be present. First, the system must have a target product. The second component is an engineered nuclear receptor that is capable of bind and activating in response to the target product but not the substrate. Finally, a library of engineered enzymes is required. If an engineered enzyme from the library is capable of producing the desired product, then the product will bind to the nuclear receptor and growth will be observed (Figure 6.1). To determine whether the survival of yeast can be linked to enzyme-activated growth, as a proof of principle, a known enzymatic pathway in the vitamin D biosynthesis pathway was used.

The biosynthetic pathway for $1,25(OH)_2D_3$ includes a three step process. First, 7-dehydrocholesterol in the skin is converted to cholecalciferol after exposure to UVB (315 nm- 280 nm) by the opening of the ring system. Next, cholecalciferol is hydroxylated at the 25-carbon by 25-hydroxylase, an enzyme in the liver, to create the product 25-hydroxyvitaminD_3. 25-hydroxyvitaminD_3 is finally hydroxylated at the 1α-carbon by 1α-hydroxylase, an enzyme in the kidneys, to create 1α,25-dihydroxyvitaminD_3, which is the biologically active compound (Figure 6.2) [8, 9].

Two different pathways of 1α,25-dihydroxyvitaminD_3 biosynthesis were tested in chemical complementation for enzyme activated growth. First, the conversion of 7-dehydrocholesterol into cholecalciferol using ultra-violet rays (UV) was investigated. As seen in Figure 6.1, if the 7-dehydrocholesterol (substrate) is converted by UV to cholecalciferol (product), the cholecalciferol should be able to bind and activate the hVDR variant, H305F/H397Y (engineered NR), allowing yeast to survive. The second targeted pathway, the conversion of the substrate cholecalciferol into 25-hydroxyvitaminD_3 by the enzyme 25-hydroxylase was investigated. If cholecalciferol (substrate) is converted into 25-hydroxyvitaminD_3 (product), then this molecule should be able to bind the wild-type hVDR (NR) and yeast should survive as a result of the enzymatic reaction.

6.2 Converting 7-dehydrocholesterol into Cholecalciferol

6.2.1 Testing hVDR and Variants with 7-dehydrocholesterol

The first step of the biosynthetic pathway is the conversion of 7-dehydrocholesterol into cholecalciferol. Exposing 7-dehydrocholesterol to ultraviolet rays (UV) causes the opening of the second ring of 7-dehydrocholesterol, producing cholecalciferol. In this

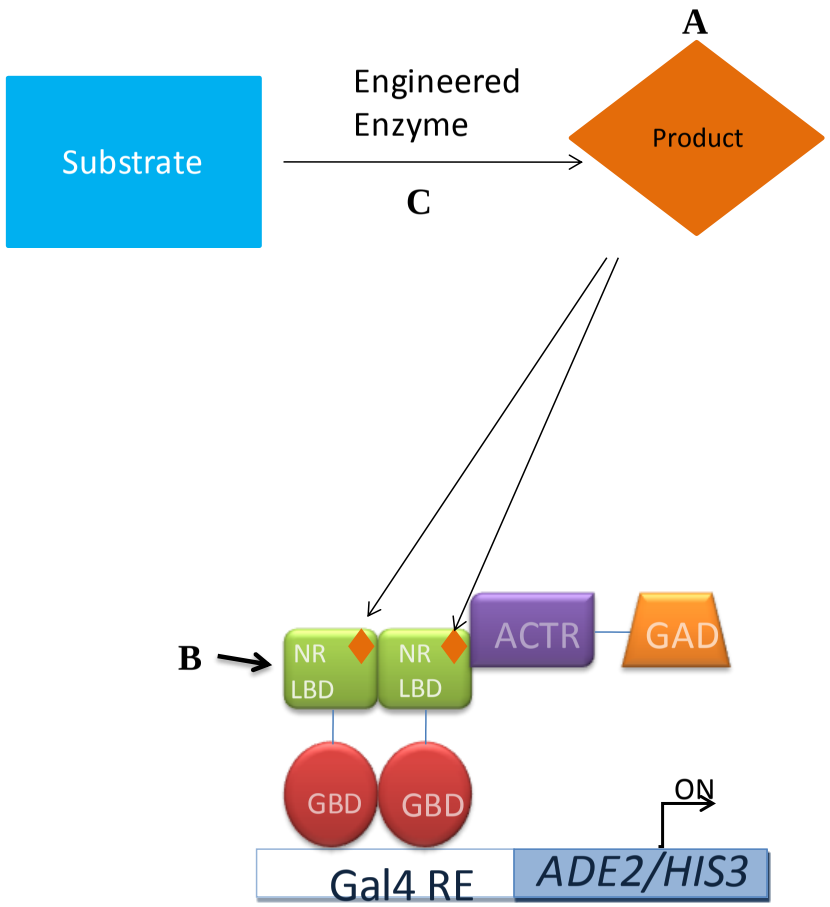

Figure 6.1: Enzyme Activated Growth Schematic. A) Target Product, B) Engineered nuclear receptor, and C) Engineered enzyme library.

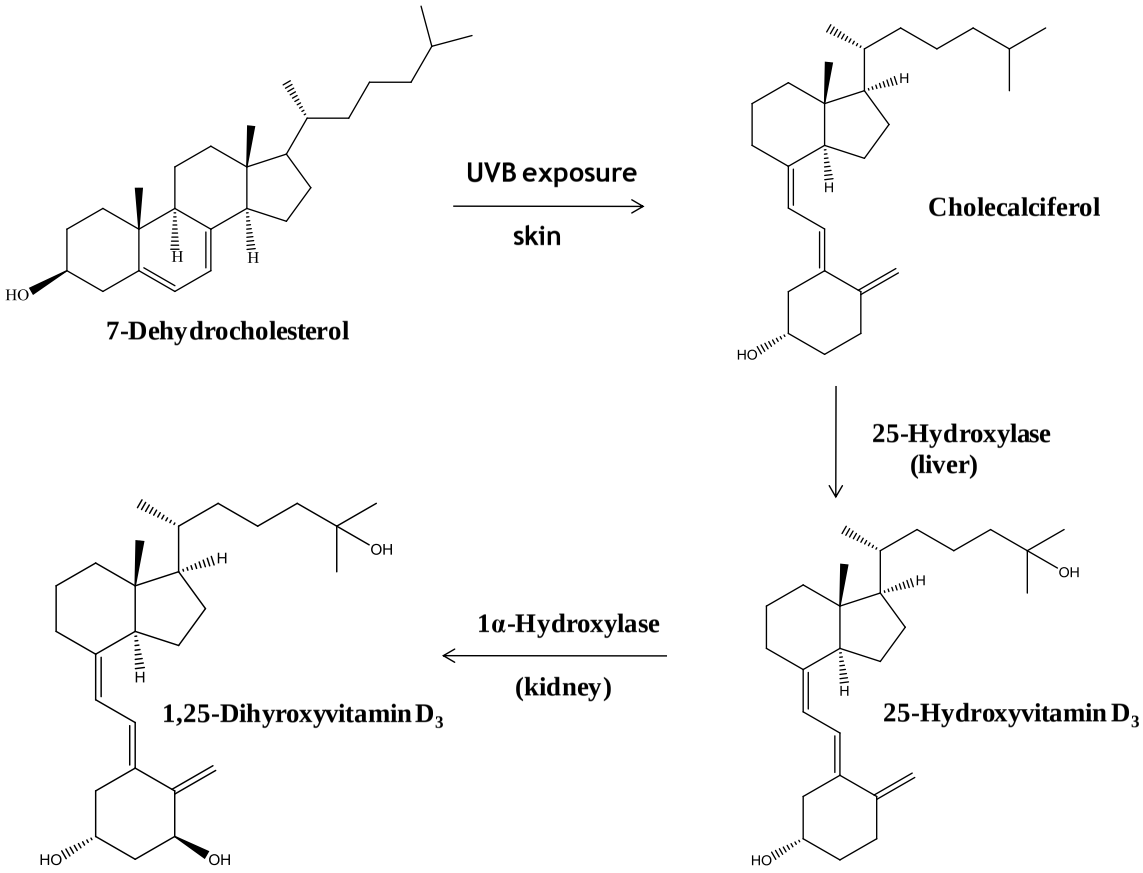

Figure 6.2: 1α,25-DihydroxyvitaminD_3 Biosynthetic Pathway

system, if the 7-dehydrocholesterol is converted into cholecalciferol, ligand-activated growth should be observed with H305F/H397Y, which has been shown to be activated by cholecalciferol, whereas wild-type hVDR should not display growth.

Before testing the conversion of 7-dehydrocholesterol to cholecalciferol, wild-type hVDR and H305F/H397Y were tested with 7-dehydrocholesterol in yeast to confirm that this ligand, which serves as the substrate (Figure 6.1) does not activate wild-type hVDR or H305F/H397Y. Wild-type hVDR , H305F/H397Y, and Gal4 (ligand-independent control) were tested in yeast with 7-dehydrocholesterol and cholecalciferol. As shown in Figure 6.3A, wild-type hVDR and H305F/H397Y did not display activation in response to 7-dehydrocholesterol. Gal4 shows growth independent of ligand. As a confirmation, wild-type hVDR is not activated by cholecalciferol, whereas H305F/H397Y shows ligand-activation at an EC_{50} of 5 μM and a 7-fold activation with cholecalciferol in adenine selective media (Figure 6.3B). Based upon these results, there should be no background growth from 7-dehydrocholesterol when testing for conversion to cholecalciferol.

6.2.2 Using UV to Convert 7-dehydrocholesterol into Cholecalciferol

6.2.2.1 Experimental Set-up

Each plasmid was transformed into yeast along with pGAD10BAACTR: Gal4, wild-type hVDR, H305F, H305Y, and H305F/H397Y. Wild-type hVDR, H305F, and H305Y were used as negative controls, due to their lack of activation with cholecalciferol. Therefore, growth should not be observed on the 7-dehydrocholesterol plates or the cholecalciferol plates. Gal4 was again used as the positive control, and should grow on all of the plates. H305F/H397Y was the protein that was being monitored for enzyme

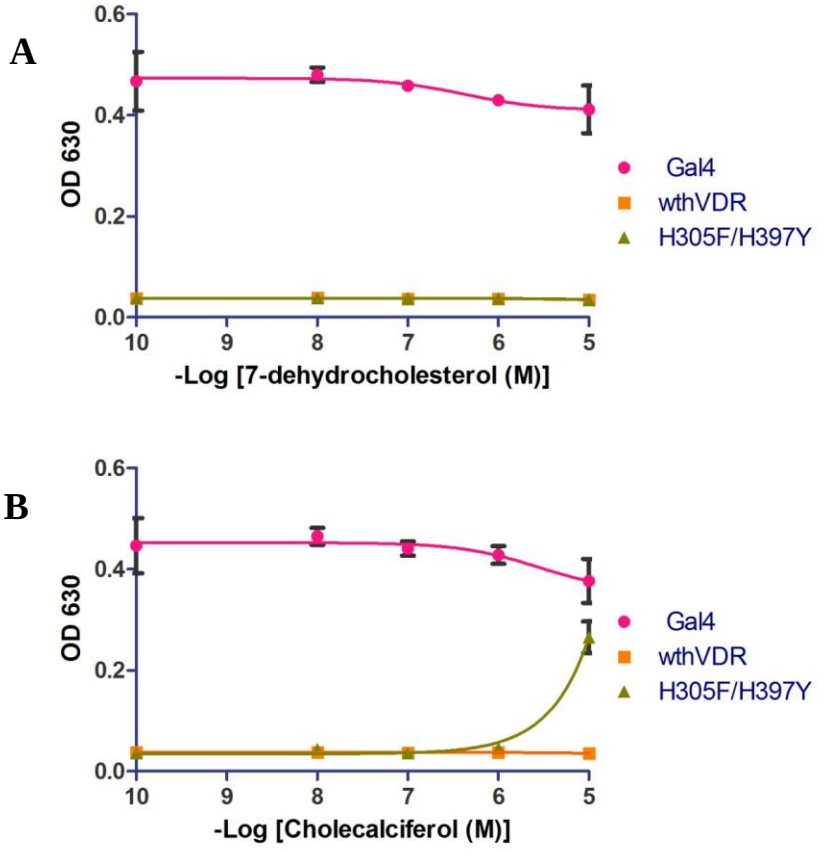

Figure 6.3: Testing 7-dehydrocholesterol and cholecalciferol in yeast.

activated growth. If 7-dehydrocholesterol (substrate) is converted into cholecalciferol (product) then yeast containing H305F/H397Y (engineered NR) should grow on the 7-dehydrocholesterol plates that are exposed to UV (represents enzyme).

Yeast were streaked onto non-selective plates (-LW), adenine selective plates without ligand (-ALW), and adenine selective plates with 10^{-5} M 7-dehydrocholesterol or 10^{-5} M cholecalciferol, as well as yeast rich media (YPD) plates. PJ694A, the yeast strain lacking any NR or enzyme, was used as a negative control and should only grow on YPD plates. Gal4, the positive control, should grow on all plates, whereas, wild-type hVDR, H305F, and H305Y should show growth on the –LW and YPD plates. H305F/H397Y and H305F/F406S, another variant capable of activating in response to cholecalciferol, should grow on –LW, YPD, and cholecalciferol plates. If 10^{-5} M 7-dehydrocholesterol is converted to at least 10^{-5} M cholecalciferol, then H305F/H397Y and H305F/F406S should grow on the UV-exposed 7-dehydrocholesterol plates.

6.2.2.2 Experimental Results

The plasmids were streaked onto two of each plate; one that was exposed to UV at 302 nM and one that was not exposed to UV. The UV plates were exposed to UV for 5 minutes. After the exposure, yeast cells containing the NR:GBD fusion and the coactivator:GAD fusion, were streaked onto these plates. The plates were incubated at 30 °C for 3 days. As shown in Figure 6.4, all sectors of the –LW plates display yeast growth except for the PJ694A sector as expected. The YPD plates displayed growth in all sectors as expected. Only the Gal4 sector of the adenine selective plate lacking ligand displayed growth as expected. Growth was observed on the cholecalciferol plates in the Gal4, H305F/H397Y, and H305F/F406S sectors as expected. No growth was observed

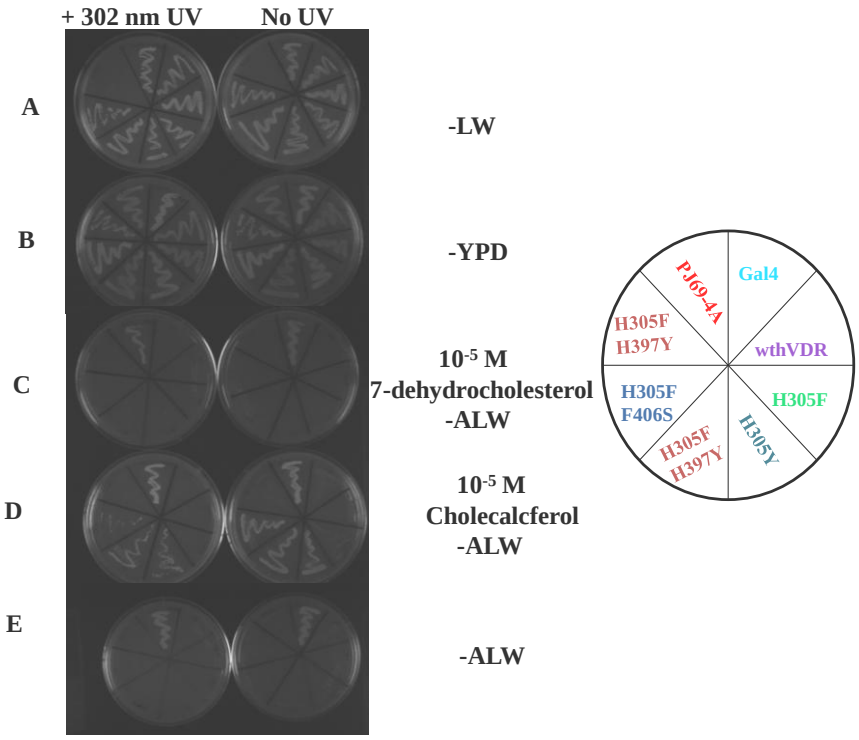

Figure 6.4: 5 minute UV exposure streaking results.

on the 7-dehydrocholesterol plates, except for the sector containing Gal4. This indicates that the conversion was not efficient enough for ligand binding and activation with the H305F/H397Y and H305F/F406S variants (Figure 6.4).

To explore whether the efficiency of the conversion of 7-dehydrocholesterol to cholecalciferol was the issue, the UV exposure time was increased to 30 minutes, and as can be seen in Figure 6.5, the increase in exposure time did not affect the results. The sector of the plate containing Gal4 displayed growth on all of the plates. However, Gal4, H305F/H397Y, and H305F/F406S were the only sectors where growth was observed on the cholecalciferol plates, as expected. As observed with 5 minute exposure, no colonies grew on the 7-dehydrocholesterol plates, except for the sector containing Gal4. Data has shown that H305F/H397Y is activated by 10^{-5} M cholecalciferol, therefore all of the 7-dehydrocholesterol would need to be converted in order to achieve ligand activation. Accordingly, the decision was made to increase the 7-dehydrocholesterol concentration to 10^{-3} M.

A new concentration of 7-dehydrocholesterol of 10^{-3} M was added to the plates. The streaking was repeated with the same plates, controls, and variants as stated above. Plates were exposed to UV for 30 minutes and then the yeast were streaked and incubated at 30 °C for 3 days. As shown in Figure 6.6, no growth was observed on the plates containing 7-dehydrocholesterol. The sector containing Gal4 should have grown on the 7-dehydrocholesterol plates; however, no growth was observed (Figure 6.6 E). This was most likely due to the precipitation of the 7-dehydrocholesterol causing a detrimental change in the media for the plates. Therefore, because the higher concentration of 7-dehydrocholesterol could not be used, the focus was shifted to the 25-hydroxylase

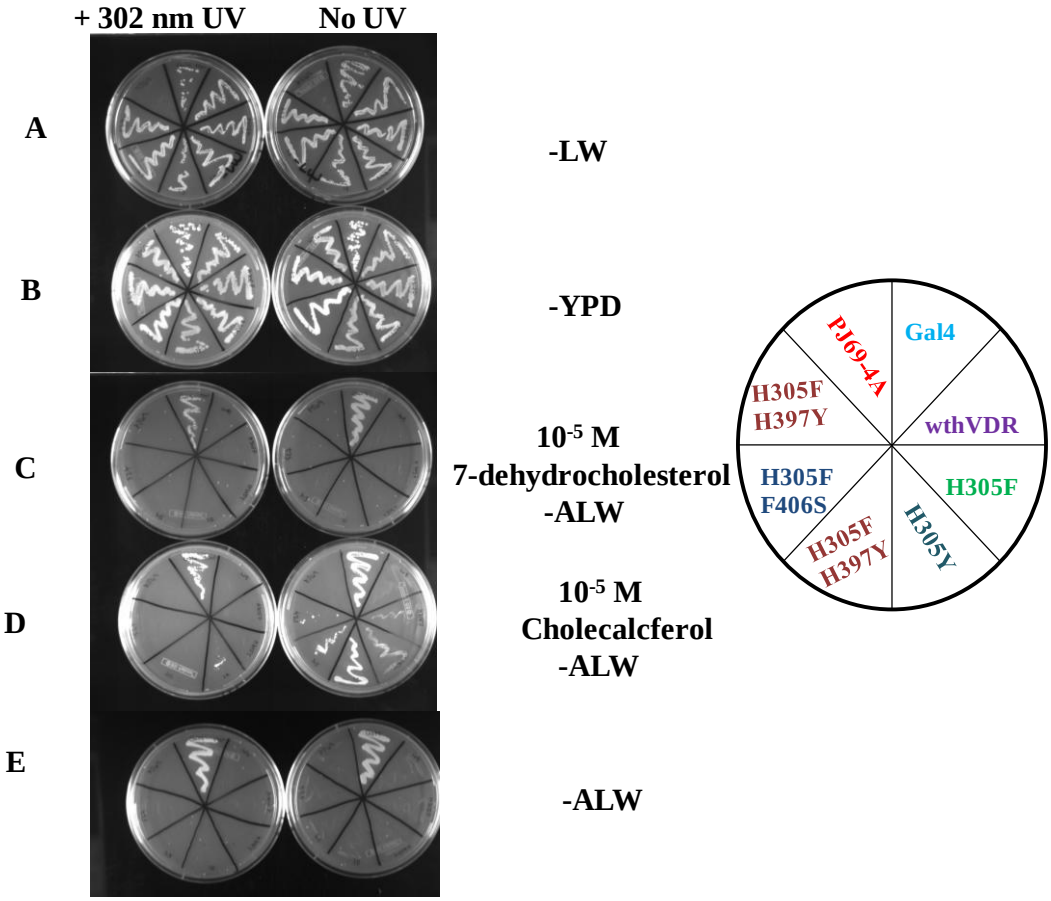

Figure 6.5: 30 minute UV exposure streaking results.

reaction in the biosynthetic pathway.

6.3 Converting Cholecalciferol to 25-hydroxyvitaminD_3

6.3.1 Introduction

25-hydroxylase (CYP2R1) is a cytochrome P450 enzyme comprised of 501 amino acids. CYP2R1 is located in the liver and responsible for converting cholecalciferol into 25-hydroxyvitaminD_3 [8, 10-12]. This enzyme has been shown to have a k_{cat} of 0.97 $\pm$ 0.05 min^{-1} as well as a K_m of 0.45 $\pm$ 0.16 µM with cholecalciferol as the substrate in yeast, which implies that the turnover rate for this enzyme is fairly slow [12].

As can be seen in Figure 6.7, if 25-hydroxylase (enzyme) can convert cholecalciferol (substrate) into 25-hydroxyvitaminD_3 (product), then the 25-hydroxyvitaminD_3 should bind and activate wild-type hVDR, inducing yeast growth in adenine selective media. The ability to use this step in the biosynthetic pathway is due to the fact that wild-type hVDR is not activated by cholecalciferol; therefore, no ligand activated growth will be observed with wild-type hVDR unless the cholecalciferol has been converted into 25-hydroxyvitaminD_3.

6.3.2 Cloning 25-hydroxylase from cDNA

25-hydroxylase is involved in the second step of the 1,25$(OH)_2D_3$ biosynthetic pathway, which takes place in the liver. The 25-hydroxylase was cloned out of liver cDNA, due to the over expression of this enzyme in the liver. The gene was isolated and amplified from liver cDNA via PCR and cloned into a yeast over expression plasmid, pUGPD, which contains a uracil selection marker and a GPD promoter. The resulting plasmid, pUGPD25-hydroxylase, was confirmed via PCR diagnostic (Figure 6.8).

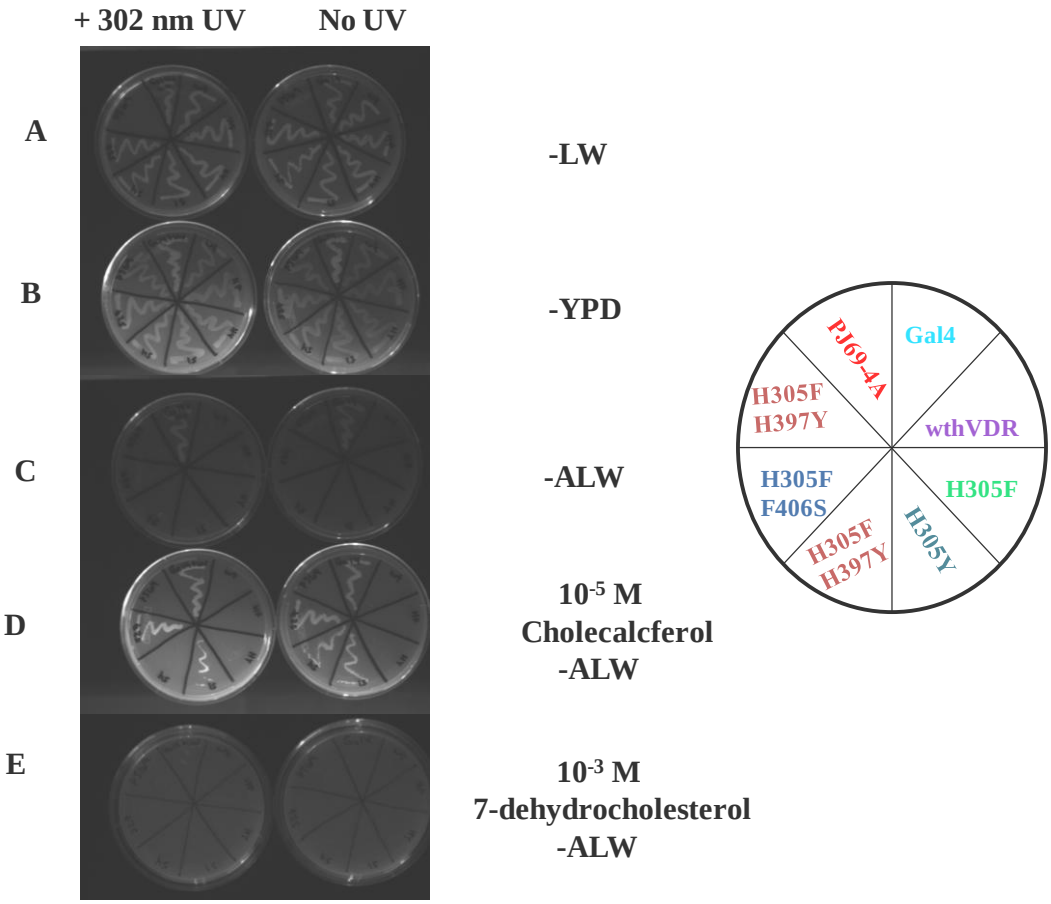

Figure 6.6: 30 minute UV exposure streaking results.

6.3.3 Testing hVDR with 25-hydroxyvitaminD_3

25-hydroxyvitaminD_3 is known to activate wild-type hVDR in mammalian cell culture. However, the ability of wild-type hVDR to be activated by this ligand in yeast had not been confirmed. In order for enzyme-activated growth to be observed, wild-type hVDR was tested for activation with 25-hydroxyvitaminD_3 in chemical complementation. Gal4 was used as the ligand-independent control. As can be seen in Figure 6.9, ligand-activated growth is observed with both wild-type hVDR and H305F/H397Y at an EC_{50} of 1 µM and 9-fold activation with 25-dihydroxyvitaminD_3. This shows that if cholecalciferol is converted into 25- hydroxyvitaminD_3 in sufficient amounts, ligand-activated growth should be observed with wild-type hVDR.

6.3.4 Experimental Design

The yeast strain BAPJ694A, which contains a *URA3* selection marker in addition to *ADE2* and *HIS3*, was used for this set of experiments. A series of plasmids was transformed into this yeast strain. Gal4 (pGBT9Gal4), wild-type hVDR (pGBDhVDR) with coactivator, H305F/H397Y (pGBDH305F/H397Y) with ACTR, and 25-hydroxylase (pUGPD25-hydroxylase) were transformed into yeast on non-selective plates, selecting for the plasmids. Wild-type hVDR and H305F/H397Y were transformed with and without 25-hydroxylase. The plasmids were then tested in liquid quantitation in adenine selective media (-AULW) with varying concentration of cholecalciferol (100 nM – 33.3 µM) to test for conversion to 25-hydroxyvitamin D_3. Gal4 was expected to display yeast growth, independent of ligand. Wild-type hVDR and ACTR, and 25-hydroxylase were negative controls and were not expected to display ligand activated growth.

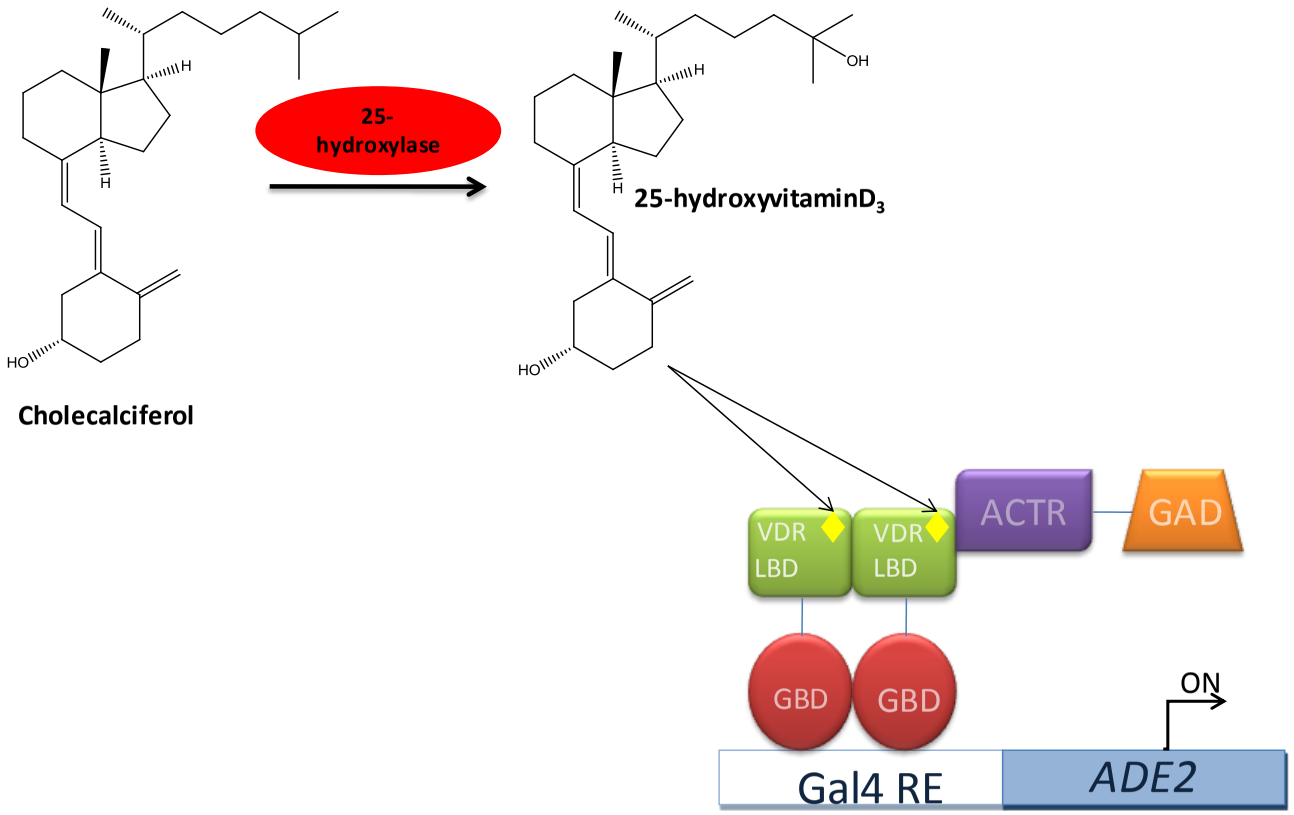

Figure 6.7: hVDR Specific Enzyme-activated growth schematic.

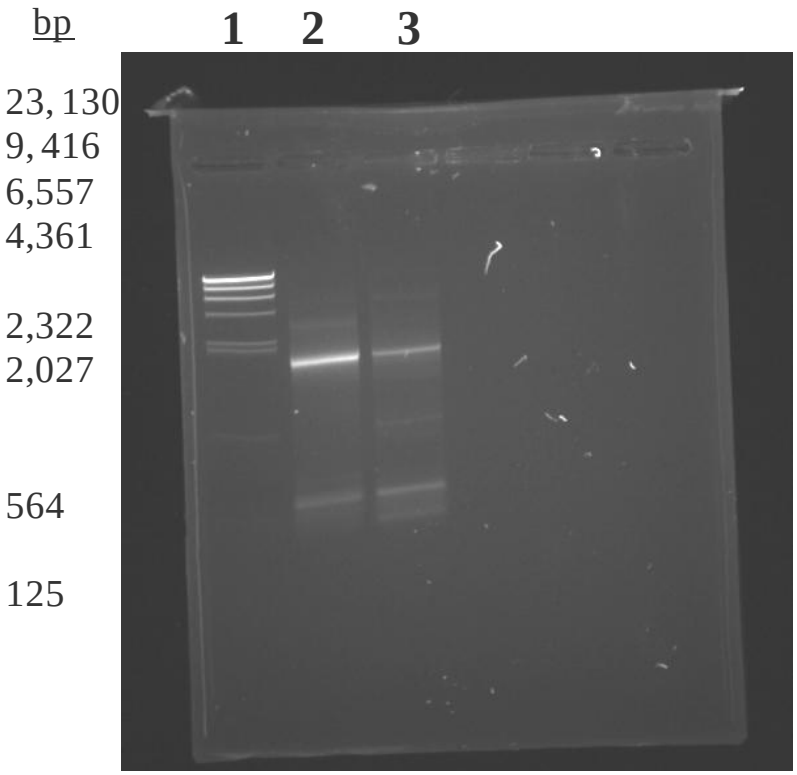

Figure 6.8: Gel of 25-hydroxylase PCR with Sequencing Primers. 1= Lambda DNA protein ladder, 2 and 3= 25-hydroxylase from 2 different PCR tubes.

H305F/H397Y and ACTR, as well as H305F/H397Y, ACTR, and 25-hydroxylase should display ligand-activated growth at 10 μM cholecalciferol in yeast. Wild-type hVDR, ACTR, and 25- hydroxylase should only display yeast growth upon the conversion of cholecalciferol to 25-hydroxyvitamin D_3 by 25-hydroxylase. The yeast were originally incubated for 48 hours and the results can be seen in Figure 6.10. At 48 hours no ligand activated growth is observed with the wild-type hVDR construct containing 25-hydroxylase. Ligand-activated growth was observed for both H305F/H397Y constructs. However, as can be seen in Figure 6.10, the H305F/H397Y construct that contains 25-hydroxylase shows ligand-activated growth with an EC_{50} of 5 μM and 8-fold activation with cholecalciferol, whereas H305F/H397Y without 25-hydroxylase displays an EC_{50} of 1 μM and 8-fold activation.

Considering these results and the slow enzyme turnover for 25-hydroxylase enzyme, the experiment was repeated but was incubated for 96 hours. The incubation time was increased in an effort to allow the enzyme more time for conversion. These results can be seen in Figure 6.11. As shown at 48 hours, the wild-type hVDR construct with 25-hydroxylase does not display ligand activated growth. The main difference is observed between the two H305F/H397Y constructs. The construct without 25-hydroxylase is constitutively active after 96 hours. However, the construct containing 25-hydroxylase is ligand-activated with an EC_{50} of 5 μM and a 7-fold activation with cholecalciferol. These results suggest that the presence of the 25-hydroxylase in the yeast is having some effect on the ligand activation of the receptor, and most likely some amount of conversion is occurring.

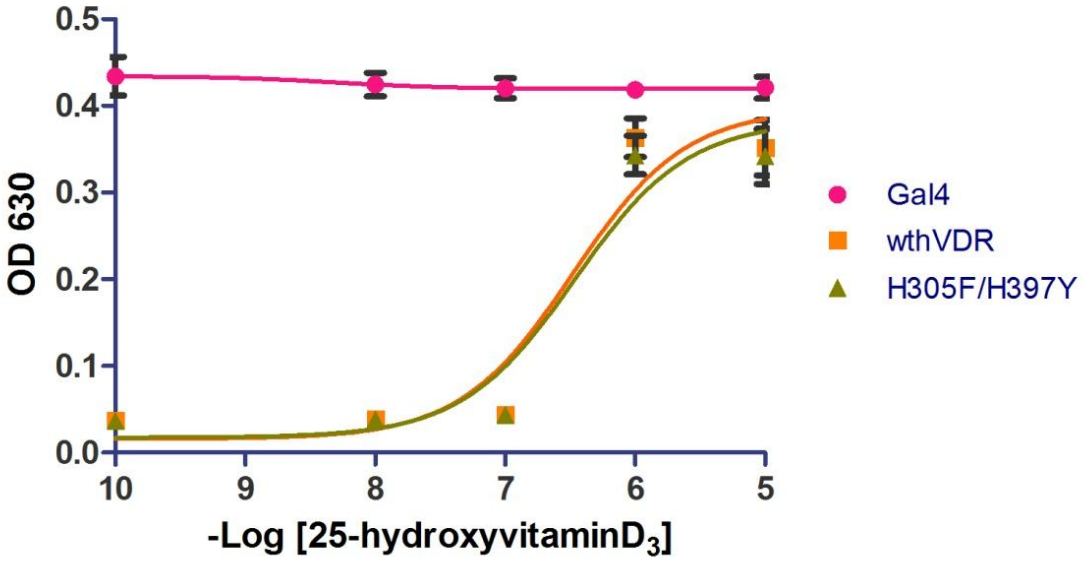

Figure 6.9: Testing wild-type hVDR and H305F/H397Y with 25-hydroxyvitaminD_3.

6.3.5 Extraction of Yeast Lysate

Based on the differences observed between the constructs with and without the 25-hydroxylase, additional information was needed regarding the effect the 25-hydroxylase has in yeast. To determine whether the conversion of cholecalciferol into 25-hydroxyvitaminD_3 via 25-hydroxylase, was occurring, a liquid quantitation assay was performed for 96 hours, and the yeast were lysed and analyzed via HPLC to determine the presence of 25-dihydroxyvitaminD_3. The organic compound was extracted from the yeast lysate, based on a protocol from Shinkyo *et al.* [12]. This protocol uses a chloroform extraction, in which an organic solution was used to allow the compounds to preferentially migrate into the desired fraction. After the extraction, HPLC experiments were performed to determine whether the reaction was taking place.

6.3.6 HPLC Results

High-performance liquid chromatography (HPLC) is a technique used to separate a mixture of compounds into the individual components of the mixture. Mixtures are placed onto a column and pushed through the column by a mobile phase (i.e. acetonitrile). The detector assigns retention times specific to each compound, serving as a method of identification.

HPLC was performed on the cell lysate to determine if any of the cholecalciferol was being converted into 25-hydroxyvitaminD_3. As can be seen in Figure 6.12, the green line represents 33.3 µM cholecalciferol, and the peak for cholecalciferol can be observed around 35 minutes. The navy blue line represents 33.3 µM 25-hydroxyvitaminD_3 with a peak around 13 minutes. The cell lysates were run on the HPLC and the spectra can be

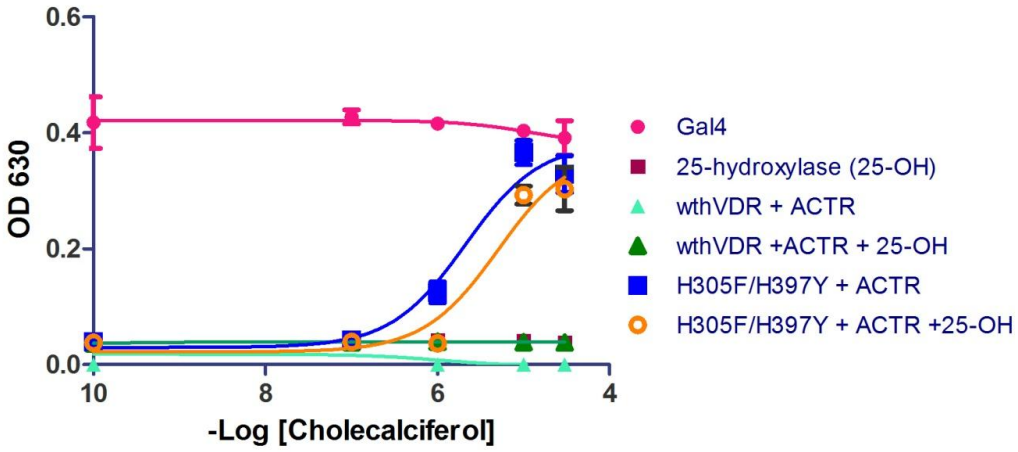

Figure 6.10: 25-hydroxylase Enzyme-activated Growth in Yeast at 48 hours.

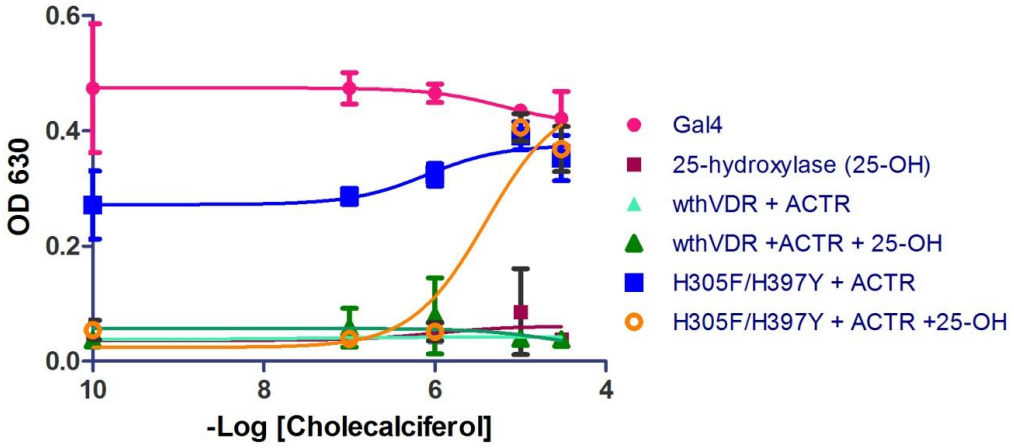

Figure 6.11: 25-hydroxylase Enzyme-activated Growth in Yeast at 96 hours.

visualized in Figure 6.12. There were two negative controls, indicating that only the nuclear receptor and coactivator were present in those yeast and therefore none of the cholecalciferol should have been converted. The black line and blue line represent wild-type hVDR and H305F/H397Y, respectively. As can be seen in Figure 6.12, the peak at 35 minutes that should represent cholecalciferol, is not observed in the controls. This implies that cholecalciferol is breaking down over the course of the 96-hour experiment. The two positive controls, which contain 25-hydroxylase in addition to the nuclear receptor and coactivator, are shown in pink and maroon. Neither a cholecalciferol peak and/or a 25- hydroxyvitaminD_3 peak is observed. This is most likely a consequence of the breakdown of the cholecalciferol and, therefore, enzyme-activated growth cannot be observed due to insufficient production of 25-hydroxyvitaminD_3.

Sequencing of the pUGPD25-hydroxylase plasmid was never confirmed, lacking confirmation of 25-hydroxylase present in the system. Despite the lack of sequencing, the presence of the enzyme in the chemical complementation does cause some effect, as can be seen in Figure 6.11. Confirmation of the presence of the correct enzyme will allow for that variable to be eliminated. Due to the lack of sequencing and HPLC data, further studies must be conducted to determine whether the enzyme is fully functional in yeast.

6.4 Summary

Based on the results seen in Figure 6.11, there is a difference in activation profiles between the nuclear receptor without 25-hydroxylase and the nuclear receptor with 25-hydroxylase. An assumption can be made that the 25-hydroxylase is having some type of effect on the yeast. However, when HPLC was performed to confirm the conversion of

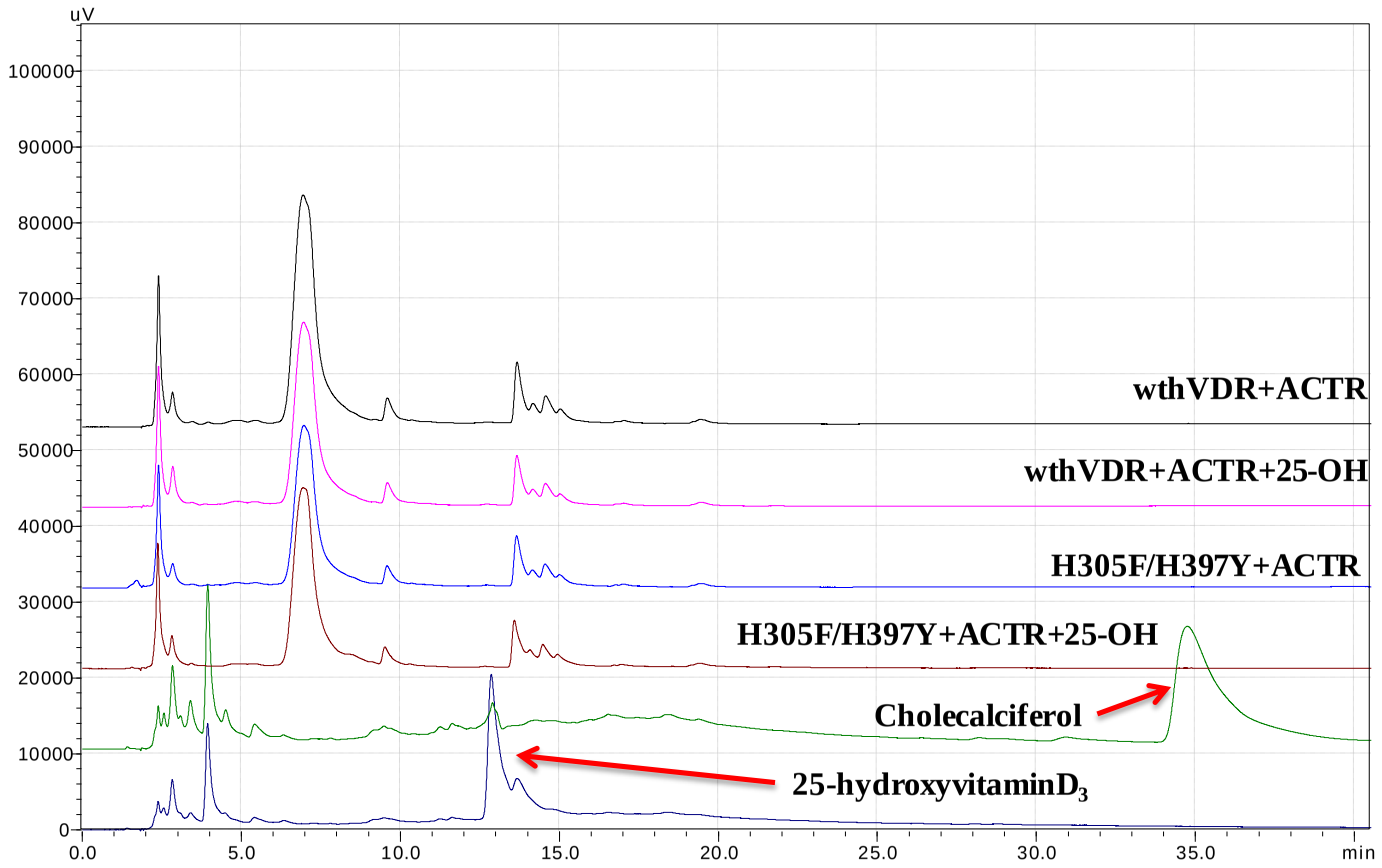

Figure 6.12: HPLC Data.

cholecalciferol into 25-hydroxyvitaminD_3, no conversion was observed due to the breakdown of the cholecalciferol over the course of the experiment.

6.5 Methods and Materials

6.5.1 Liquid Quantitation Assays

Plasmids were tested in liquid quantitation assays in 96-well plates with media lacking adenine, leucine, and tryptophan (SC-ALW), with or without 7-dehydrocholesterol, cholecalciferol, and 25-hydroxyvitaminD_3 at concentrations ranging from 0.01 µM - 10 µM. A 4:1 ratio of media (SC-ALW): cells (yeast resuspended in water) were aliquoted into 96-well plates. Plates were incubated at 30°C, with shaking at 170 rpm. Optical density (OD) readings at 630 nm were recorded at 0, 24, and 48 hours as a measure of growth density.

6.5.2 Streaking for 7-dehydrocholesterol Testing

Yeast plates were made with or without 7-dehydrocholesterol and cholecalciferol. The plates that were made were SC-LW, SC-ALW without ligand, SC-ALW with 7-dehydrocholesterol or cholecalciferol, and YPD plates. Plates were exposed to ultraviolet rays at 302 nm for 5 minutes or 30 minutes. The yeast were then streaked onto the plates and incubated at 30 °C for 2 days.

6.5.3 Cloning 25-hydroxylase from cDNA

The 25-hydroxylase gene (CYP2R1) was cloned from human liver cDNA with PCR using the following primers: 5’- aaggaaaaaa gcggccgc atgtggaagctttggagag-3’ and 5’- tcc ccgcgg tcagcgtctttcagcac-3’. The underlined sequences denote *NotI* and *SacII* restriction sites, respectively. The 25-hydroxylase gene and yeast expression vector (pUGPD), were digested, ligated, and transformed into Z-competent *XL-1 Blue E.coli* cells (Zymo, USA).

Fusing the 25-hydroxylase to the Gal response elements created the pUGPD25-hydroxylase plasmid, which contains a uracil marker.

6.5.4 Testing Conversion of Cholecalciferol in Liquid Quantitation

The following plasmids were tested in liquid quantitation with cholecalciferol: pGBT9, containing Gal4, pGBDhVDR and ACTR and25-hydroxylase, pGBDH305F/H397Y and ACTR and 25-hydroxylase, pGBDhVDR and ACTR, pGBDH305F/H397Y and ACTR, and pUGPD25-hydroxylase. Plasmids were tested in liquid quantitation assays in 96-well plates with media lacking adenine, leucine, uracil, and tryptophan (SC-AULW), with or without cholecalciferol at concentrations ranging from 0.001 µM - 33 µM.

6.5.5 Lysing Yeast

The 96-well plates were centrifuged at 1500 rpm for 15 minutes and the supernatant was removed. Lyticase (10 mg/mL) was added to the plates and the pellets were resuspended. The plates were then incubated for 30 minutes at 37 °C. Triton (0.1% v/v) was added to the plates and incubated at room temperature for 10 minutes. The plates were then centrifuged at 1500 rpm for 15 minutes. The supernatant was transferred to a new 96-well plate.

6.5.6 Extraction and HPLC

The supernatant for each sample was extracted with four volumes of chloroform/methanol (3:1 v/v). The organic phase was recovered by evaporating the chloroform/methanol off of the sample. The residue was dissolved in acetonitrile and applied to the HPLC. HPLC was performed on a YMC-Pack ODS-AM column (4.6 x 300 mm). A linear gradient of 70-100% acetonitrile for 15 minutes followed by 100%

acetonitrile for 25 minutes was the mobile phase. The column temperature was maintained at 40 °C.

6.6 References

1. Azizi B, Chang EI, Doyle DF: **Chemical complementation: small-molecule-based genetic selection in yeast**. *Biochem Biophys Res Commun* 2003, **306**(3):774-780.

2. Schwimmer LJ, Rohatgi P, Azizi B, Seley KL, Doyle DF: **Creation and discovery of ligand-receptor pairs for transcriptional control with small molecules**. *Proc Natl Acad Sci U S A* 2004, **101**(41):14707-14712.

3. Taylor JL, Rohatgi P, Spencer HT, Doyle DF, Azizi B: **Characterization of a molecular switch system that regulates gene expression in mammalian cells through a small molecule**. *BMC Biotechnol*, **10**:12.

4. Baker K, Bleczinski C, Lin H, Salazar-Jimenez G, Sengupta D, Krane S, Cornish VW: **Chemical complementation: a reaction-independent genetic assay for enzyme catalysis**. *Proc Natl Acad Sci U S A* 2002, **99**(26):16537-16542.

5. Elander RP: **Industrial production of beta-lactam antibiotics**. *Appl Microbiol Biotechnol* 2003, **61**(5-6):385-392.

6. Demain AL, Elander RP: **The beta-lactam antibiotics: past, present, and future**. *Antonie Van Leeuwenhoek* 1999, **75**(1-2):5-19.

7. Wegman MA, Janssen MHA, van Rantwijk F, Sheldon RA: **Towards biocatalytic synthesis of beta-lactam antibiotics**. *Adv Synth Catal* 2001, **343**(6-7):559-576.

8. Cheng JB, Levine MA, Bell NH, Mangelsdorf DJ, Russell DW: **Genetic evidence that the human CYP2R1 enzyme is a key vitamin D 25-hydroxylase**. *Proc Natl Acad Sci U S A* 2004, **101**(20):7711-7715.

9. Jones G, Strugnell SA, DeLuca HF: **Current understanding of the molecular actions of vitamin D**. *Physiol Rev* 1998, **78**(4):1193-1231.

10. Cheng JB, Motola DL, Mangelsdorf DJ, Russell DW: **De-orphanization of cytochrome P450 2R1 - A microsomal vitamin D 25-hydroxylase**. *J Biol Chem* 2003, **278**(39):38084-38093.

11. Strushkevich N, Usanov SA, Plotnikov AN, Jones G, Park HW: **Structural analysis of CYP2R1 in complex with vitamin D3**. *J Mol Biol* 2008, **380**(1):95-106.

12. Shinkyo R, Sakaki T, Kamakura M, Ohta M, Inouye K: **Metabolism of vitamin D by human microsomal CYP2R1**. *Biochem Biophys Res Commun* 2004, **324**(1):451-457.